

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:46:57 ; Search time 7316.32 Seconds

(without alignments)
4710.839 Million cell updates/sec

Title: US-09-826-581-5

Perfect score: 1647

Sequence: 1 ttggtctggggctggccaca.....acacagcctctagctctc 1647

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_un:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlgc_inv:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	DB ID	Description
------------	-------	--------------	-------	-------------

1	1647	100.0	1647	6	AX281582	AX281582 Sequence
2	1590	96.5	2290	6	HS2429977	AX2429977 Homo sapi
3	1453	88.2	2115	6	AX099802	AX099802 Sequence
4	1453	88.2	2115	9	AF214519	AF214519 Homo sapi
5	1447	87.9	2109	6	AX099776	AX099776 Sequence
6	1172.8	71.2	2022	6	AX099804	AX099804 Sequence
7	1140.4	69.2	1873	4	AF214520	AF214520 Sus scrofa
8	1140.4	69.2	1873	6	AX099800	AX099800 Sequence
9	1134.4	68.9	1867	6	AX099774	AX099774 Sequence
10	1134.4	68.9	1867	6	AX099774	AX099774 Sequence
11	414.8	25.2	1328	10	RNMPKGM	X95578 R. norvegicus
12	407.2	24.7	989	6	AX281579	U42413 Rattus norv
13	407.2	24.7	196554	2	AC073128	AC073128 Sequence
14	396.8	24.1	1623	10	AF035635	AF035635 Mus muscu
15	395.4	24.0	3132	10	BC015283	BC015283 Mus muscu
16	395.2	24.0	206854	3	AC009974	AC009974 Homo sapi
17	393.4	23.9	1578	9	HS042413	U42412 Human 5'-AM
18	393.4	23.9	1677	9	BC000358	BC000358 Homo sapi
19	366.6	22.3	1167	9	AB025580	AB025580 Homo sapi
20	366.6	22.3	1435	6	AR139104	AR139104 Sequence
21	366.6	22.3	2194	9	AF087875	AF087875 Homo sapi
22	366.6	22.3	2203	9	BC020540	BC020540 Homo sapi
23	366.6	22.3	2223	9	AK001887	AK001887 Homo sapi
24	363.4	22.1	2062	9	HS2429976	AF094764 Drosophila
25	300.8	18.3	2082	3	AF094764	AF094764 Drosophila
26	299.2	18.2	4666	3	AY070541	AY070541 Drosophila
27	278.8	16.9	1014	6	AX281581	AX281581 Sequence
28	278.8	16.9	152129	2	AC027416	AC027416 Homo sapi
29	278.8	16.9	196554	2	AF073128	AF073128 Homo sapi
30	273.6	16.6	5888	4	AF214521	AF214521 Sus scrofa
31	267.4	16.2	227724	2	AF336381	AF336381 Mus muscu
32	198.2	12.0	152129	2	AC027416	AC027416 Homo sapi
33	168.2	10.2	1722	6	AX281580	AX281580 Sequence
34	131	8.0	3497	3	AF094763	AF094763 Drosophila
35	127.2	7.7	80069	2	AC019671	AC019671 Drosophila
36	127.2	7.7	173634	3	AC009344	AC009344 Drosophila
37	127.2	7.7	195868	3	AC008308	AC008308 Drosophila
38	127.2	7.7	230266	3	AE003733	AE003733 Drosophila
39	91.4	5.5	106	11	G67375	G67375 A898 Human
40	84.8	5.1	151	6	AX305717	AX305717 Sequence
41	83	5.0	142903	2	AL627254	AL627254 Danio rer
42	78.4	4.8	110000	2	CEY111B2.1	Continuation (2 of
43	78.4	4.8	110000	3	CEY111B2.2	Continuation (3 of
44	78.4	4.8	257702	3	CEY111B2A	AI132904 Caenorhab
45	75	4.6	14411	4	AF329081	AF329081 Bos tauru

ALIGNMENTS

RESULT 1
AX281582
LOCUS AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (sites)
Andersson, L., Luthman, H. and Marklund, S.
Variants of the human amp activated protein kinase gamma 3 subunit
Patent: WO 0177305-A 5 18-Oct-2001;
JOURNAL
Arenis AB (SE)
Location/Qualifiers
1..1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
20..1489
/note="unnamed protein product"
/codon_start=1

CDS

file
copy

Sequence search results

/protein_id="CAD10589.1"
/db_xref="GI:1608834"
/translation="MERGELHRLRRTPSMSSLGSSEHOEHSFLEOENSSSPSPAYTS
SSERIRGRRAKALRMTRKROKSEGEERGGRRSPAPAEENSCIEPTKPTPIAOA
DPAGVTPPGMDLPSDCTASAGSSDDVELATERPAENSCIEPTKPTPIAOA
LSRAPPKRLKDELKRPQAITMRMOEHCTIDAAATSKLVEFTMLEIKRFEFA
LVANGRAAPLMDSKOSFVGMATITDFILVLRHYSPLVOJIELOHKTIEWRELY
LQGFKPLVSLPNDLSFEAVYTLIKNRILHPLDPSGNVLIHLTKRLKLEHIF
GSLPRSPFLRTIADLIGIFRDLAVLEATLTADLFEDRVSALPVNBEQGV
VGLSRSDVILAAOQTYNHLMDSVGALRORCTLEGIVLSQPHESLGEVIDRIARE
OVRHVLVDETHLGYVSLSDILOALVSPAGIDALGA"

BASE COUNT 346 a 502 c 462 g 337 t
ORIGIN

Query Match 100.0%; Score 1647; DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ttgtcttggtggtgacacatgagccgggtgagacgactgacagacccttc 60
1 ttgtcttggtggtgacacatgagccgggtgagacgactgacagacccttc 60
61 ctggagcagccttggtggttctgagcatcaagaagtgtcttctagagcaagaacag 120
61 ctggagcagccttggtggttctgagcatcaagaagtgtcttctagagcaagaacag 120
121 cagctcatggtccatcacacagctgtgacagcagctcagaagaatcctgaggaacggag 180
121 cagctcatggtccatcacacagctgtgacagcagctcagaagaatcctgaggaacggag 180
181 ggcgaagaaccttgagatgagacagagcagaagtctggtgaggaagaaggagcacaagtca 240
181 ggcgaagaaccttgagatgagacagagcagaagtctggtgaggaagaaggagcacaagtca 240
241 ggggaaagtgctcccggtcagaagcagctgctgagtcacaacgggtcagagcagacattccc 300
241 ggggaaagtgctcccggtcagaagcagctgctgagtcacaacgggtcagagcagacattccc 300
301 caagaacacaccttggtcagatgacatcctgacgggtgagcactccacacacagaggtg 360
301 caagaacacaccttggtcagatgacatcctgacgggtgagcactccacacacagaggtg 360
361 ggaactgctccctctgactgttaagcctcagctgacagctcagacagagatgagtgtga 420
361 ggaactgctccctctgactgttaagcctcagctgacagctcagacagagatgagtgtga 420
421 gctggcaccagagtgctcagacagcagagcctgagagtgtagctataagaagcctgctgga 480
421 gctggcaccagagtgctcagacagcagagcctgagagtgtagctataagaagcctgctgga 480
481 agagaagcctgacctgctgctgctcccgcaagcccatctcccaagctgagctgtgagta 540
481 agagaagcctgacctgctgctgctcccgcaagcccatctcccaagctgagctgtgagta 540
541 cgaactgagaaaccggcgcccgagatctatagcgtctatgacagagacacctccta 600
541 cgaactgagaaaccggcgcccgagatctatagcgtctatgacagagacacctccta 600
601 cgaatgcaatgcaactaactcaagctagctatcttgcacacatgctgtgagatcaaga 660
601 cgaatgcaatgcaactaactcaagctagctatcttgcacacatgctgtgagatcaaga 660
661 ggcctcttctgctgtgagcacaagctgtgagcagacccctctatgagcagcaaga 720
661 ggcctcttctgctgtgagcacaagctgtgagcagacccctctatgagcagcaaga 720
721 gcaagagcttggtgagagtgatgacacatgacagcttcatctgtgtgtgactgacta 780
721 gcaagagcttggtgagagtgatgacacatgacagcttcatctgtgtgtgactgacta 780
781 caggtcccccctggtccagatctatgagatgacaataagaattgagacctgagga 840
|||||

781 CAGTCCGCCCTGCTCAGATCTATGAGATTGAACAACATTAAGATTGAGACCTGGAGGA 840
841 gatctactgcaagagctgtcttaagccttctgctcactctctctaatgataagcctgt 900
841 gatctactgcaagagctgtcttaagccttctgctcactctctctaatgataagcctgt 900
901 tgaagctgttaacacctcaacaaacggatgacatgctcctgctgctcttgagaccgt 960
901 tgaagctgtgtacacctctatcaagaacggatgacatgctcctgctgctcttgagaccgt 960
961 gtaaggaacgtatccacatctcacaacaacgctgctcaagttcctgacatctt 1020
961 gtaaggaacgtatccacatctcacaacaacgctgctcaagttcctgacatctt 1020
1021 tgggtccctgctgctcccgccctctctctacacgacatcccaagattggacatgg 1080
1021 tgggtccctgctgctcccgccctctctctacacgacatcccaagattggacatgg 1080
1081 cactctcagagactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1140
1081 cactctcagagactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1140
1141 cttgtgagacggcgtgtgtctgactgctgctgctgctgctgctgctgctgctgctg 1200
1141 cttgtgagacggcgtgtgtctgactgctgctgctgctgctgctgctgctgctgctg 1200
1201 cctctatcccgcttggatgagatcactgctgctgctgctgctgctgctgctgctgctg 1260
1201 cctctatcccgcttggatgagatcactgctgctgctgctgctgctgctgctgctgctg 1260
1261 catgagtggtgagaaagcctcagagcagagacatctgctgagagagagctcttctctg 1320
1261 catgagtggtgagaaagcctcagagcagagacatctgctgagagagagctcttctctg 1320
1321 ccagcccaacgagagctgtggtgagagtgatcagacagatgctggtgagagagctcag 1380
1321 ccagcccaacgagagctgtggtgagagtgatcagacagatgctggtgagagagctcag 1380
1381 gctgtgctagtgagaaagcctcagacatctctggtgctgctgctcctctcagacatct 1440
1381 gctgtgctagtgagaaagcctcagacatctctggtgctgctgctcctctcagacatct 1440
1441 tcaagcactggtgtcagcctgctgctgctgctgctgctgctgctgctgctgctgctg 1500
1441 tcaagcactggtgtcagcctgctgctgctgctgctgctgctgctgctgctgctgctg 1500
1501 tctcaatcccaagacacactgacacactggaagccaatgaaagagagcagaaactcagc 1560
1501 tctcaatcccaagacacactgacacactggaagccaatgaaagagagcagaaactcagc 1560
1561 ctcatcttcccccaccccatcttggctgctgctgctgctgctgctgctgctgctgctg 1620
1561 ctcatcttcccccaccccatcttggctgctgctgctgctgctgctgctgctgctgctg 1620
1621 ggcacatgacacacagcctttagcttc 1647
1621 ggcacatgacacacagcctttagcttc 1647

RESULT 2
HSA249977
LOCUS HSA249977 2290 bp mRNA linear PRI 07-APR-2000
DEFINITION Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
(AMPK gamma 3 gene).
ACCESSION AJ249977
VERSION AJ249977.1 GI:6688200
KEYWORDS AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2290)
Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.

TITLE	Characterization of AMP-activated protein kinase gamma-subunit isoforms and their role in AMP binding
JOURNAL	Biochem. J. 346 Pt 3, 659-669 (2000)
MEDLINE	20164049
REFERENCE	2 (bases 1 to 2290)
AUTHORS	Carling D.
TITLE	Direct Submission
JOURNAL	Submitted (12-Oct-1999) Carling D., Cellular Stress Group, MRC

Biochem J. 346 Pt 3, 659-669 (2000)

```
gene      /organism="Homo sapiens"  
          /db_xref="taxon:9606"  
          22..1500  
CDS       /gene="AMPK gamma 3"  
          22..1500
```

```

/evidence=experimental
/product="AMP-activated protein kinase gamma 3 subunit"
/protein_id="CAB65117.1"
/db_xref="GI:6688201"
/translation="MEPELEHALLRTPTSSSLGSEHDEMSFLDQENSSMPSRAVITS
SSEIRIGKRRAKALRMTROKSEVEGEPFGQCEGRSPRAASTGIEAEPFTTPIAQQ
DPAVGTPIPTGMDCLPSDCTASAGSSSTDVELNTEPATAMSCIEGLEIEEPALCL
LSPAPFPEKLGMDLELRKPGAQIIMRLEFETHCYCDAMATSKLVIPTMLIKKAFEF
LVAGVRAAPLSPDLSKOSQSFVGLMTITFPIILVHRYNPSQVIOIETIBQKTEMRREIY
LOGCEKPLVASISPDLSLEAVYLTILKNHILPLVADPAGVNLITPKRLKLTHTFI
GLLSPRPSTLITDIDIGICFGRDLAVYLENAPLITLALDIEVDKRVSLAPVNLKCGQV
VGLSKRPDVIHLAAQOTYTNHLDMSVGEALNRKIDTCEGLVSCQPHSESGEIVDIRAE
QVHLVLVDENQHLLGVASLSDILQALVPLSPAGIDPSPGPEKTI"

```

Query Match	96.5%	Score 1590:	DB 9:	Length 2290:
Best Local Similarity	99.3%	Pred. No. 0:		
Matches 1607:	Conservative	0:	Mismatches	10: Indels
				1: Gaps
QY 1	ttgtgtctggggccttgcgcacatcagcagccccgggctctgagcacgcactctgcagagacccttc	60		
Db 3	TTGGCTCTGGGGCTGGCCCATGTGAGAACCCCGGGCTGGAGACGACGACTGTGGCAGACCCTTC	62		
QY 61	ctgagcagcgtcttgggggttcttgagcatcaagagatgagcttcctcaaggcaagaacaag	120		
Db 63	CTGGAGCAGCCTTGGGGGTTCTGAACATCAAGATGAGCTTCTTAGAGCAAGAAAG	122		
QY 121	cagctcatlpgccatccatccagcctgtgacagcagctcagaagaatcctgtggaacgag	180		
Db 123	CAGCTCATGGGCATCCACACGCTGTGACCGACGAGCTCAAGAAAGAAATCGTGGGAACGGAG	182		
QY 181	ggccaaagcctcttgagatlgacaagcagaatctgttgaagaaagggggccaaccagtcca	240		
Db 183	GGCCAAAGCCTTGAAGATGAGCAAGGCAAGATCGGTGAGGAAGGGGAGCCACCAAGGTCA	242		
QY 241	gggggaaggtcccccgttccaggcagcctgtcctagtlccaccgggctctgagggccacaattccc	300		
Db 243	GGGGGAAGGTCCCGCGTCCAGGCGAGCTGTGAGTCCACCGCGGCTGGAGGCCACATTGCC	302		
QY 301	caagaccacaccctctgtcctaagctgataccttcggcggggttgggcaatccacaacagggtg	360		
Db 303	CAGACACACACCTTTGGCTCAAGCTGAATCTGCGGGGGTGGCACTCCACCAACAGGGGTG	362		
QY 361	ggagctgcctccctctgactgtacagcttcaagctctgcaagctccacagacagatgtatgttga	420		
Db 363	GGACTGCCTCCCTCTGACTGTACAGCCTCAGCTGACAGCTCCAGCAAGATGATGTGA	422		
QY 421	gcttgcacaggaagtccacagccacagagggccttggagatgtagctatagaaggctgtgtga	480		
Db 423	GCTTGCAACGGAGTTCCTCCAGCCACAGAGGCTGGGAGTGTAGCTAGTAAGGCTCTGTGGA	482		

QY	481	agagaagcctgcctcgtgtcctcgtccgcgcagccccatcttccaaagctgtgtcgtggaatga	540
Db	483	AGAGAAGCCTGCCTCCTGTGCTGTCCCGCAGGCCCCCAATTTCCCAAGCTGTGGCGGATTA	542
QY	541	cgaactcgtgaaaccgcgcgcgaagatcatactgcttcaatcgaagaaaccgcgtca	600
Db	543	CGAACTCGGGAACCCGGCGCCAGATCTTCATTCGCTTATCGAGAGCACACCTGCTTA	602
QY	601	cgatgcgatgycgaactagcttccaaagctagtcattcttcgaacacatgcttggataccaaga	660
Db	603	CGATGCCAATGGGCMACTACTCTCAAGCTAGTACATCTTTCGACACCATGCTGGATCAAGAA	662
QY	661	ggcctctcttctgtctcgtgtgagcaacggcttbgcggagagccccctatggaacgaaga	720
Db	663	GGCCTCTTCTTGTCTCTGTGTGGCCAAAGGTGTGGGGGAGCCCTCTATGGAGCAGAAAGAA	722
QY	721	gcagaagcttctgtgaggaatgctgcgaacatacgaacttcaatccctgtgtcgtacgtacta	780
Db	723	GCAGAGCTTGTGTGGGGATGTGACCAATCACATGACTTTCATCTGTGTGTGATGCGTACTTA	782
QY	781	caggtccccccctggtgcagatctctatggaattgaaacaataagaattggaaccttgagga	840
Db	783	CAGGTCCCCCCTGGTGCAGATCTTATGAGATTGAACAACATMAATTGAGACCTCGAGGGA	842
QY	841	gattactactgaaagctgtgcttcaagctctgtgtccatctctccaaatgaatagcctgt	900
Db	843	GATCTACCTGCAAGGCTGTCTCAAGCCTGTGTCTTCATCTCTCTTAATGATTAAGCTGT	902
QY	901	tgaagctgtctaacacctcatcaagaacgcgataccatcgctcgtctgtctcttgaaaccggt	960
Db	903	TGAAGCTGTCTTAACACCCCTCATCAAGAACCCGATATCCCTGCTCCGTCTTGACCCGCT	962
QY	961	gtcagaagcagctactccaacatctcccaacaagaacgcctgtcctaagtctcttcaatctt	1020
Db	963	GTCAGAGCAACTACTCCACATCTCTCACACACAAAGCCTGCTCAATTCCTCTCACATCTT	1022
QY	1021	tggttccctgtgcgccggccctctctctctcaacgcacataccaagaatttggacacg	1080
Db	1023	TGGTTCCTCTGCTGCCGGCGGCTCTCTTCCTCAACCGACTATCAAGATTGGGACTGG	1082
QY	1081	caacatctcgagaacttggtctgtgtgtcgtggagaaagaaaccaatctcgtactgcacgtgaac	1140
Db	1083	CACATCTCGGAACCTTGGCTGTGTGTGGTGGAGAACGACCCATCTCTACGCACTGTGAAT	1142
QY	1141	cttgttgaaaccggtgtgtgtctgactgcctgtgtcacaagaaatgtgtcagaatcgttgg	1200
Db	1143	CTTGTGTGAACCGGGGTGTGTCTGTGACAGCTGTGTGTCACGAATGTGGTCAAGTGTGG	1202
QY	1201	cctcaattcccggtcttgatgtgatttaacctggtgtgccagcaaaactaaacaacctgga	1260
Db	1203	CCTCAATTCCTCGCTTATGTATGTATGTATGATCGTGGCTGCCACGAAACCTTAACACACTCGGA	1262
QY	1261	catgagtgtggagaaagccctctgaggtgcgaagaaacaatgctctggaagggtctctctctg	1320
Db	1263	CATGAGTGTGGGAAGACCCCTGAGAGGAAGACATATGTCTGAGAGGAGTCTTCTCTG	1322
QY	1321	ccagccccacgaaggtcttgggggaagtgtatcgacaggaattgtctggagagcaagtatacacag	1380
Db	1323	CCAGCCCCACGAGAGCTTTGGGGGAAGTATGATCGACAGATGATGCTCGGGAGCAGGTACACAG	1382
QY	1381	gctgtgtctagtgaacgaagaccagaatctctctgtggcgtgtgtcctcctccgaaatctct	1440
Db	1383	GCTGTGTCTAGTGTGACGAGACCAAGCATCTTGTGGCGTGTGTCTCTCTCTCCGAATCTCT	1442
QY	1441	tcaaggcaactgtgtctcaagccctgcgtgcaatcgaaatgtgcctctccgggctctgagaagaatctgag	1500
Db	1443	TCAAGCACTGTGTCTCAAGCCCTCTGTGGCATTCATCTCCCTCGGGGCTGTGAAAGATCTGAG	1501
QY	1501	tcctcaatctccaagccacccctgcacacctctggaagccaatgaagaaggaactcggagaactcagc	1560
Db	1502	TCCTCAATCTCCAAACCCACTCTGCACACCTCTGGAAGCCCATTAAGGGAATCTGGAAATCTGAC	1561


```
/protein_id="CAD10589.1"
/db_xref="GI:1660834"
/translation="MEPELEHALRTPSMVSLGSEHOENSLFEQENSSSPSPAYTS
SSERIRKRRRAKALRMTRKOKSVEGEPGEGEPGPAESGLETFPKTTPLOA
DPAGVTPPTGMDCLPSDCTSAAGSSTDEVELATEPATPAEAGEEGLLEERPLAC
LSPOAPEPKLGMDELKPCGAOYIMRMOHCTYDMATSSKIYIDYHLEIKAPFA
LVANGVAAPALMDSKOSFYGMLTITDFILVLRYYRSPLOYLEYEONHETWREY
LOCFRPLVSISSPNDSLFEAVYTLINKRHRPLVLPVSGNVHLILTRKLFLEHF
GSLRPSRFLRTIODIGICTFRDLAVLEETAP1LTALDIFDVRVRSALPVEGCOV
VGLYSRFDVYLLAAQOYTNLMDSVGEALRORTLCLEGVLSQPHESLGEVIDRIARE
QVHRVLVDETQHLGLVSVLSLIDLOALVISPAGIDALGA"
```

BASE COUNT 346 a 502 c 462 g 337 t

ORIGIN

Query Match	100.0%	Score 1647	DB 6	Length 1647
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1647	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	ctggctcgggagctggagccatgagccggagctggagcagcagctgagcagcccttc	60		
DB 1	TTGGCTCTGGGGCTGGCCACATGAGACCCGGGCTGGAGCAGCACTGGCAGACCTTTC	60		
QY 61	ctggagcagccttgggggtcttgagca tcaagagatgagcttccttagagcaagaacag	120		
DB 61	CTGAGACGACGCTTGGGGTCTTGAGCATCAAGAGATGAGCTCTTACAGCAAGAAACAG	120		
QY 121	cagctcatgagccatcaccagcttgagcagcagctcagaaagaatccgtgggaaacggag	180		
DB 121	CAGCTCATGGCCATGACCACCTTGACACCACTGCAAGAAAGAAATCCGTGGAAACGGAG	180		
QY 181	ggccaaagccttggatgagcaagagcagaaagctggtgaggaaggagccacaggtca	240		
DB 181	GGCCAAAGCCTTGAGATGAGCAAGAGCAAGAGTGGTGGAGAAAGGAGCCACCAAGTCA	240		
QY 241	gggggaaaggtcccccgtccagagccagctctgagctccaccgggctggaggccaatccc	300		
DB 241	GGGGGAAGGTCCCGGTCGAGGCCAGCTCTGAGTCCACCGGGCTGGAGGCCACATCCC	300		
QY 301	caagaccacaccccttggctcgaagctgactcgtcgggggtgggagctccacaacaggtg	360		
DB 301	CAAGACCACACCTTGGCTCAAGCTGATCTCCGGGGGTGGCACTCCACCAACAGGGTG	360		
QY 361	ggaactgcccctctgactgactagcagctcagctcagagctccagacagatgatgtga	420		
DB 361	GGACTGCTCCCTTGACTGACTGACGCTCAGCTGCGAGGCTCCAGCAGCAATATGTGGA	420		
QY 421	gtcgggacagagatcccaagccacagagcctggaagtgtgagctagaagggctctgga	480		
DB 421	GCTGGCACCGACATTTCCACACAGAGGCTTGAGCTGAGAGGGCTTGAGAGGGCTCTGGA	480		
QY 481	agagagagcctgcccctgtccgtgcccgagggcccaattcccaagctgggctggagta	540		
DB 481	AGAGAGGCGCTGGCCTGTGCTGTCCCGGAGGCCCATTTTCCAAAGCTGGGCTGGATGA	540		
QY 541	cgaactgcggaaccccgagcccaagatctacatgctcctcaatgacagagacacccgtga	600		
DB 541	CGAACTGCGGAACCCGGGCCCAAGATCTAATGCGCTTCATCAGAGACACACCTGCTA	600		
QY 601	cgaatgcataggcaactagctcaagatagctcctcgacaacca tgcgtgagatacaaga	660		
DB 601	CGATTGCCATGGCACTAGCTCCAAAGCTAGCTATTCGACACCATCTCGAGATCAAGAA	660		
QY 661	ggcctctcttgccttgtaggcaacagtg tgcggcagccctctatggaagcaagaaga	720		
DB 661	GGCCTTCTTGTGCTGTGGTGGCAAGGTGTGGGGCAGGCCCTCTATGGGACAGCAAGAA	720		
QY 721	ggagagcttctggggagatgctgaaccaatcactcactgctggtgctgcatcgtacta	780		
DB 721	GCGAGACTTGTGGGATCTGACATCACTTCACTGAGTGTGTCATCGCTACTA	780		
QY 781	caggtcccccgtgcacatctatgagatctgaacaacaatgattgagacctggaggga	840		
DB 781		840		

DB 781	CAGTCCCCCTGGTCCAGATCTATGAGATTGAACAACATAGATTGAGACTGGACGA	840
QY 841	gattactctcaagagcgtgtcaagcctctggtctcattctccctaataatagccgtt	900
DB 841	GATTACTCTCAAGAGCTGTCTTCAAGCTCTGTGCTCCATCTCTCTAATATAGCTGTT	900
QY 901	tgaagctgtctaacctctcaacaaacggagatccalcgctcgtctgtcttcaaccgt	960
DB 901	TGAAGCTGTCTAACCTCTATCAAGAAACGGAGATCATCGCTGCTGTCTTACCCGGT	960
QY 961	gtcaggaagagctactccaaactctcaacacaaagcgtctcaagtctcgcacatct	1020
DB 961	GTCAGGCAACGTACTCCACATCTCTCAACACAAACGCTGCTCAAGTCTCTGACATCTT	1020
QY 1021	tggctccctgctgcccggccctcctcctcaacgcacatccaaagattgggactcg	1080
DB 1021	TGGTCCCTCTGCCCCGGCCCTTCTCTACCCACATATCCAAATTTGGGCATCGG	1080
QY 1081	caaatcccgagacttgctgtggtgctgagacagcaaccaatccctgactgagacat	1140
DB 1081	CAAATTCGAGACTTGGCTGTGCTGGAGACGACCACTCTGACTGCACTGGACAT	1140
QY 1141	ctctgtgagccgctgtgtctgactgctgtgtgtcaacgaatgtggtcaggtctgg	1200
DB 1141	CTTGTGACCGGCGCTGTCTCTCACTGCTGTGTCAACGAATGTGTAGGTCTGTGG	1200
QY 1201	ccctcatctccgcttctgattgattcactgtgtgtcccaacacctaaccacttga	1260
DB 1201	CCCTCATCTCCGCTTGTGATGTGATTCACCTGGCTGCCCAACAACTTACACCACTTGA	1260
QY 1261	catgaatgtgtgggaaagcccttgaggagaggaactatgtctggaaggagctctctctg	1320
DB 1261	CATGAATGTGTGGGAAGCCCTTGAGGACAGACACATCTGTGAGGAGTCTTCCCTG	1320
QY 1321	ccagccccaagagctctggggaaagtatcgacaagatctgtctcgggagcaggtacag	1380
DB 1321	CCAGCCCAAGAGCTTGGGGAAAGTATCGACAGAGATTTGCTCGGAGACAGTACACAG	1380
QY 1381	gctgtgtgtagtgagcagaaccagacatctcttggtgctgtgtccctctccgacatct	1440
DB 1381	GCTGTGTGTAGTGTGAGAGAACCCAGCATCTTGTGGCGGTGCTCCCTCCGACATCTT	1440
QY 1441	tcaagcactgtgtctagcctctggtgcatatgctcccgaggctctggaaagatctgg	1500
DB 1441	TCAAGCACTGTGCTTAGCCTGTGCTGCAATGATCCCTGGGGCTTGAAAGATCTGAG	1500
QY 1501	tccatcatcccaagccaacttgacaccccttgaaagccaatgaaggaaactggaactcagc	1560
DB 1501	TCCTCATTCCAAGCCACTGCAACCTGGAAGCCCAATGAAGGAACCTGAAACTCAAC	1560
QY 1561	cttcatcttcccccaaccctcatttctgtgttcaagatagatcagtaggtctgccttg	1620
DB 1561	CTTCAATCTTCCCCCAACCCCATTTCTGTGTTCAAGTATGATTCAGGTAGGCTGCCCTG	1620
QY 1621	ggcatgacacagcagcctctagcttc 1647	
DB 1621	GGCCATGACACCAAGCTCTTAGTCTTC 1647	

RESULT 2
HSA249977
LOCUS HSA249977
DEFINITION Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit (AMPK gamma 3 gene).
ACCESSION AJ249977
VERSION AJ249977.1 GI:6688200
KEYWORDS AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2290)
Citeung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.

QY	1561	ctcaatcttccccccccccatttgcgttccagctatgattcaggttgcgc	1618
Db	1562	ctttcatcttccccccaccccatcttctgcttcacgtatcatcagctttctttgacgtc	1619
RESULT	3		
LOCUS	AX099802	2115 bp	DNA linear PAT 02-APR-2001
DEFINITION	Sequence 29 from Patent WO0120003.		
ACCESSION	AX099802		
VERSION	AX099802.1		
KEYWORDS	GI:13538836		
SOURCE	human.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2115) Anderson, L., Looft, C., Kalm, E., Milan, D., Rodic, A., Rogel-Gallard, C., Iannucci, N., Gellin, J., Le Roy, P. and Chardon, P.		
TITLE	Variants of the gamma chain of ampk, dna sequences encoding the same, and uses thereof		
JOURNAL	Patent: WO 0120003-A 29 22-MAR-2001; INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ; Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)		
FEATURES	Location/Qualifiers		
source	1..2115		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
CDS	1..1395		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC35801.1"		
	/db_xref="GI:13538837"		
	/translation="MSFLRDSNSMSPSPAVTSSEKIRGRKRAKLRWTRKSVSEEG EPPOGQSPRSRPSTESTGLBATEPKTPPLAQADPAVGVPTPTWDCLPDQCTASAG STDDVLELATETPAETMEWCELEGLLEERALCLSPQAPPKLMDDELKRPQAOIYM RFMEHCTCYDAMATTSKLVIFDTMLELTKAEFALVANGVAAPLMDSKOSFVOMLT1 TDFLVLHRYRSP.LVOIYELEQHKITWBEIYLQCFKPLVSPDSLFEAYYTLDI KNRIHRLPVLDPVSGNLHLITHRKRLPLHIFCSLLPRPSFLRTTLDGIGTFRDL AVLETPAPLITLALDI FVDHRSALPVNPEGGVGLSFRDVIHLAODQRYNHLDMSV GEALRORTLCLEGLVLSQPHESLGEVIDRLAREVHRLVLYDETQHLGLVYSLDILO ALVLSPADIGALGA"		
BASE COUNT	460 a 622 c 562 g 471 t		
ORIGIN			
Query Match	88.2%: Score 1453; DB 6; Length 2115;		
Best Local Similarity	97.3%: Pred. No. 0;		
Matches 1501:	Conservative 0; Mismatches 35; Indels 7; Gaps 2;		
QY	95	atgagctctcctagagcaagaagaacagcagctcatgagcatcaccagctgtgacagcagc	154
Db	1	ATGAGCTTCTCTTAGAGCAAGAAGAACAGCAGCTCATGGCCATCACAGCTGTGACCAAGCAGC	60
QY	155	tcaagaagaatccatgaggaagacgagagggccaaagccttgagatgagcaagcagaagtgc	214
Db	61	TCACAAAGATCCGTGGGAAACCGAGGGCCCAAAGCCTTGAGATGACAAAGCAAGATCG	120
QY	215	gtlgaagaaggaggaagcaccagctcagggggaagaagctcccggtccagggcagcctgag	274
Db	121	GTGAGAGGAAGGGAGGCCACACGCTCAGGGGGAAGTCCCGCGTCCAGGCCAAGCTGCTGAG	180
QY	275	tccagcgggctggagggccacatctccccaagaacacacaccccttgctcaagctgacctgc	334
Db	181	TCACACGGGCTGGAGGCCACATTCCTCCCAAGACCAACACCTTGCTCAAGCTGATCTCGCC	240
QY	335	gggggtgggcaatccacaacaaggggtgggaatgacctccctctgactgtlaacagcttaagct	394
Db	241	GGCGTGGGCACTCCACCAACAGGCTGGAGCTCCTCCCTGTACCTGTAAAGCCTTCAGCT	300
QY	395	gcaaggtccagcaagatgtagtggagactggtgccaaggaagctccagccaagaagcctgg	454

D	301	GCAGGCTCCACACAGATGATGTGGACTTGGCCACGAGATTCCACAGCCACTGG	360
O	455	gagtgtagcctagaaggccctctcgtgaagaagaagccctgcctctgtcctgtcccgaggcc	514
D	361	GAGTGTGAGCTAGAAAGCCTCTCTGTGAAGAAGAGCCTTGCCCTGTGCTGTGCCCGAGGCC	420
O	515	ccattcccaagctcgtggtctggatlgacgaactcgtcggaaaaccggcgccagatctacatg	574
D	421	CCATTGCCAAGCTGGGCTGGATGAGCAACTCGCGAAACCCGGGCCCGCATGTACATG	480
O	575	cgcttcacgaaggagcaacctctgaagatgcacatgcaactcagctccaagttagtcac	634
D	481	GCCTTCATGCAAGGAGCAACCTCTGATGATGCCATGGCAACTGACCTCAAGCTACTCATTC	540
O	635	ctcgacacacatgcctggagatccaagaagccctctctgtcctgtgtgccaagctgtgcg	694
D	541	TTTCGACACCATGCTGTGAGATCAAGAAGCCCTTTTGGCTGTGCTGTGGCAACGCTTGCGG	600
O	695	gcagccctctatctggacacgaagaacagaagctctgtgggaatgcctgaacatctgac	754
D	601	GCAGCCCTCTCATGTGGAGACACCAAGAACAAAGCTTTGTGGGATCTCAACCTACCTGAC	660
O	755	ttcaatcctcgtgcgtcatctgactacaagctcccccctgtcctcagatcatatgaattgaa	814
D	661	TTTCATCTCGGTGCTGATTCGCTACTACAGTCCCTCCCTGTGCTCAATCTATGAAATTGAA	720
O	815	caacaataagatltgagaccctggaaggaagatctacatctgcaagctgtctcaagcctctgtc	874
D	721	CAACATTAAGATTGAGACCTGTGAGGAGATCTACTCTCAAGCGCTGCTCAACGCTTGCTC	780
O	875	tcacatctctctaaatgaatccgttttgaagctgtctacaacctatcaagaaccgagtc	934
D	781	TTCATCTCTCTAATGATACCTGTTTGAAGCTGTCTACACCTCATTAAGAACCGGATC	840
O	935	catcgctgcctgtctctcttgaccggcgtgtcaaggaactcctcacatcctcacacacaa	994
D	841	CATGGCGCTGCTGTTGTTGACCCGGTGTGAGGAACTACTCCATCATCTCACAACAAAA	900
O	995	cgctctgcctcaagtctccgacatcttgttccctcgtcgtgcccgccctctctctctac	1054
D	901	GCCTGTCTCAAGTTCGTCGACATCTTTGGTTCCTCTCTCCGCCCTCTCTCTCTAC	960
O	1055	cgcactatccaagaattcttggagcatccgacatctccgaactctgtcctgtgtgtctgagaca	1114
D	961	CGCACTATCCAAAGATTGGGCGATCGGCACATTCCAGACTTGGCTGTGCTGTGGAGACA	1020
O	1115	gcaccatctctgaactgacatctgacatctcttctgtgaacggcggtgtgtctcgaactcgtg	1174
D	1021	GCACCCATCTCGATCGCAGCGACATCTTTGTGTGACCCGGCGGTGCTCGCACTCCCTGTG	1080
O	1175	gtcaagaatctgtctgaagtcgtgggcccctatctccgcttctgattgtatctacactgtct	1234
D	1081	GTCACACGAATGTGCTACGTCGTGGGCGCTCTATTCCCGCTTTGATGTATTCACCTGGCT	1140
O	1235	gccacgaacaaactiacaaaccctctgacatgagtgatgggaagaccctctgaagcagaagaca	1294
D	1141	GCCCGAGAAACCTTACAACCACTGTGACATGTGTGGAGAAAGCCCTGTAGCGACAGAGACA	1200
O	1295	ctatgctctgaaagaaatctctctctccgacggccccaagagaagcttgggggaagatgtatgac	1354
D	1201	CTATGTCTGAGGAGATCTCTTCTCTGTCCAGGCCCAAGAGACTTGTGGGAAGTATGTATGAC	1260
O	1355	aggaatctcgggaagagattacacagcctgtgtctagtgtgacgtgacacgaacatctctctg	1414
D	1261	AGGATTTCTCGGAGAGAGTACACAGCGTGGGTGTAAGTGACAGAGACCAGCATCTCTTG	1320
O	1415	ggcggtgtctccctctccgaactctcttaagagacgtgtgtcctcaagcctgtctgacatctgat	1474
D	1321	GGCTGTATCTCTCTTCGACATCTCTTAGAGAGCTGGTGTCTACCCCTGCTGGCATCTGAT	1380
O	1475	ggcctcggggctctgaagaatctgtacctcaatcccaagcgaactctgcacactctggaagc	1534
D	1381	GCCTCGGGGCTCTGAAATATCTGATGCTCTCAATCTCCAAAGCAACTCTGCACA -CTGGAGC	1439

OY	1535	caalgaaggaagcctcagaggaacacagcctcaatcctcccccaccccacattgctgltcag	1594
DB	1440	CAATGAAGCAATTTGACA-----ACAGCTTCAATTTCCTCCCAACCCCAATTTGCTGGTTCAG	1493
OY	1595	ctatgattcaggttagagctctcgcctcgtggacatgacacacagcct	1637
DB	1494	CTATGATTCAGGCTTCTTCAGGCTTCCAAATAATGGCTTTGGCT	1536

RESULT	4
AF214519	
LOCUS	2115 bp mRNA linear PRI 03-JUN-2000
DEFINITION	Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
ACCESSION	mRNA, complete cds.
VERSION	AF214519
KEYWORDS	AF214519.1 GI:8215681
SOURCE	human.

REFERENCE
AUTHORS

TITLE	A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle
JOURNAL	Science 288 (5469), 1248-1251 (2000)
DOI	10.1126/science.10818001
PUBMED	10818001

TITLE *Indirect and direct selection for carcass traits in sheep*
JOURNAL *Animal Breeding and Genetics*
Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish University of Agricultural Sciences, BMC box 597, Uppsala 751 24, Sweden

BASE COUNT	FEATURES	source	location/Qualifiers
460	a	gene	1..2115
622	c	CDS	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="2" /map="2p" /tissue-type="skeletal muscle"
562	g		1..2115
471	t		/gene="PRKAG3" 1..1395 /gene="PRKAG3" /note="PRKAG3" /codon_start=1 /product="AMP-activated protein kinase gamma subunit" /protein_id="AAF73987.1" /db_xref="GI:8215682" /translation="MSFLDENSSSWPSPAVYSSSRIRGKKRAKALRMVTRKQSVEEG EPFGGEGPRSRPTAESTGLENTFPRTTFLAADPAGVGTPTPGWDLCSIDCTASAG SSTDVDELATERPATEAMECELEGLEENRPALCLSPQAPFKLGWDELRSQAGQIYM RMQSHCTCDANATSKVLEIFDTMLEIKKAFALVANGRAALIMDSKQSPGMITI TDFILPLHYRYSPLVQIYEIOHKIETMRFEIYLQGCPEKPLVSIQNSDLFEAVYTLI KRIIRHLRVPLVSGNVNLHILTKRLLRFLRHTIGESLLRPSFLYRTIDLGIGTRDL AVLEETAPLITALDIFVDRRVSAALRVNVECGGVGLYKSEVDVILHLAQQTYGHLMSV GEALQRFLTCLEGVLTSCOPHESLGEVDIARIAREQVHRLVLDETOHLGLVSLDILQ ALVLSPAIDALGA"

Query Match	88.2%	Score 1453	DB 9	Length 2115
Best Local Similarity	97.3%	Pred. No. 0		
Matches 1501; Conservative	0	Mismatches 35	Indels 7	Gaps 2

OY	95	atagatttccttagaagaagaaaaacgacgctctatagccatcaacagctctgacagcagc	1354
Db	1	ATAGGCTTCCTTAGAGCAAGAAAAACAGCACACTTCATGGCCATCAACGAGCTGTACCAGAGC	60
OY	155	tcaagaagaatccgtctgtagaaacggagggccaaagcctctgagatgtgacaaggcagaatctg	214
Db	61	TCAGAAAGAAATCCGTGGAAACGGAGGCGCAAAAGCCTTGACATGGACACAGGCACAAAGTGC	1220
OY	215	gttagagaaaggagagccacacaggtccaagggggaagtctcccgctccagggccaagctgtgag	274
Db	121	GTGTAGGAAGAGGGGAGGACACACAGCTAGGGGGGAAGTCCCGGTGCAGGCGCAATGCTGAG	180
OY	275	tccacggagcctgtagagccacatctcccaagaagcaaacctctgtgcttaagctgatacccgcc	334
Db	181	TCACACGGGCTGGAGGCGCACATTCCTCCCAAGACACACCCTTGCTCAAGCTGATCTGGCC	240
OY	335	ggggtgtgagcactccacacaaggtgtggaactgtcctccctctgactgtatcaagcctcaact	394
Db	241	GGGGTGTGGACATCCACCAAAAGGGTGGGACGTCCCTCTGTGACTGTACAGCTTCAGCT	300
OY	395	gcagagctccagacagatgtatgttagagctgtgcccaggaattccagccacaaggagcctgtg	454
Db	301	GCAGGCTCCAGCACAGATGATGTGGAGCTGGCCACGGAGTTCCAGGACACAGAGGCTTGG	360
OY	455	ggagtgtgagctagaagagcctgtctggaagaagagcctgtccctgtgtcctgtctcccgacagcc	514
Db	361	GAGGTGTAGCTAAGAAAGGCTCTCTGGAAAGAGAGGCTGCGCTGCTGCTCCCGCAGGCC	420
OY	515	ccattctccaaagctgtgctgtggaatgtgacgaactgtggaagaaocggagcgccagatctatag	574
Db	421	GCATTTCCCAAGCTGGCTGGGATGTACGAACTGGGGAAACCCGGCGCCAGATCTACATG	480
OY	575	cgcttcatagtcaggaagacacctgtactagatgtccaatgtgcaactagatctcaagctagctatc	634
Db	481	CGCTTTCATGTGAGAGACACACCTGCTACGATGCCATGTGCACATGTAGCTCAAGCTATGCTATC	540
OY	635	ttcgcacacacatgcttggaatacaagaagagcctctctgtcctgtgtgtgcccaggtgtgctgg	694
Db	541	TTTCGACACCATGTCTGTGAGATCAAAAGAGCCTTCTTTTCTCTGTGGGCGCAACGGTGTGGCG	600
OY	695	gcagccctctatcggaagcaagaagcagagctttgtgtggaatgtctgacacatcaactgac	754
Db	601	GCAGCCCTCTATGTGGAGACGACGAAAGCAAGAGCTTTGTGGGATGCTGACATCACTGAC	660
OY	755	ttcaatcctgtgtgtgcgcatcgctataaggttcccccctgtgtccaagatctatgtagatgtgaa	814
Db	661	TTTCATCTGTGTGTGCTGCATCGCTACTACAGGTCCCGCTGTGTCAAGATCTATGAGATTGAA	720
OY	815	caacataagatltgagacttggagggagatctacctgtcaagagctgtctcaagcctctgtgtc	874
Db	721	CAACATAGATTGTGAGACTGTGAGGAGGAGATCTACTCTGCAAGCGCTTCAAGCCTCTGTGTC	780
OY	875	tccatctctctatatgtatagcctgttttgaagctgtctcaacccctcatcaagaacggatc	934
Db	781	TTCACTCTCTCTAATGTATACCTCTTTTGAAGCTGTCTACACCCTCATCAAGAAACGGATC	840
OY	935	catgcctgtcctgtctctctgacccgggtgtcagagaaagtaactcaaaactctcaacaacaa	994
Db	841	CATGCGCTGCTGTTTGTGACCCGGGTGTACGGCAAGCTACTCCATCTCTCAACACAA	900
OY	995	cgcctgtctcaagttcctgtcaacatcttggttccctgtctgcccggcgccctctctctctac	1054
Db	901	CGCTGCTCAAGTTCCTGCACTCTTTGGTTCCTCTCTCCCGGCGCTTCCTCTTAC	960
OY	1055	cgacatctcaagatltggagcatcggaacatctccgaactgtgctgtgtgtctgtgagaca	1114
Db	961	CGCAGTATCCAAATTTGGGCATCGGCACATTCGACGAGATTGGCTGTGTGTGGAGCA	1020
OY	1115	gcaacccatctgactgcaactgtgacaactcttgtgagacggcggtgtctctgcaactgtctgt	1174
Db	1021	GCACCCATCTGTGACTGTACACTGTGACATCTTTGTGTGACCCGGCGTGTCTGTGCACTGCTGTG	1080
OY	1175	gtcaacgaatgtgtgaagttcgttgggcctctatctccgcgcttltgatactaccctgtgct	1234

|||||
Db 1081 GTCACGATGTGGTCAGGTCTGGGCTCTATTCCGCTTTGATGTGATTCACCTGGCT 1140
OY 1235 gccacgaacactacaacaccctgacatgagtgtagagagagccctgagacagagaca 1294
Db 1141 GCCCAGCAACCTACAAACACCTGAGACTGAGTGGAGAGAACCCCTGAGCAAGACACA 1200
OY 1295 ctatgctcggagggagagctcttctctcgcagcccccagagagcttgagggaatgacgac 1354
Db 1201 CTATGCTGTGAGGAGTCCCTTCTGCCACGCCACAGACCTGGGGGAGATGATGCAC 1280
OY 1355 agaatctcgcagagacagatcacagagctgtagtagtagagagagaccacatctctg 1414
Db 1261 AGATTGCTCGGAGCAGAGTACACAGCTGTGCTAGTGAGCAGACCCACACATCTCTG 1320
OY 1415 ggcgtggtctcctctccgacatctcctcaggaactgtagtgcagccctgctgcatcgat 1474
Db 1321 GGCCTGCTCTCCCTCTCCGACATCTTCAGGACACTGGTGGCTCAGCCCTGCTGGATCGAT 1380
OY 1475 ggcctcggggccttgagagagatcttgagctcccaatcccaagccacctgacaccttgaa 1534
Db 1381 GCCCTCGGGCTGTAGAGATCTGATCTCAATCCCAACCCCAACTGCACAC-CTGGAAGC 1439
OY 1535 caatgaaagggaactggaactcagcctcatctcccccaccccaatctgctggttcag 1594
Db 1440 CAATGAAGGAATTGAGA-----ACAGCTTCATTTCCCAACCCCAATTGGCTGTTTCA 1493
OY 1595 ctatgtagtagtagtagctgctgacctgggacgaacacagcct 1637
Db 1494 CTATGATTTCAGGCTTCTTCAGGCTTCMAAATGCCCTTTCCT 1536

RESULT 5
AX099776 2109 bp DNA linear PAT 02-APR-2001
LOCUS AX099776
DEFINITION Sequence 3 from Patent WO0120003.
ACCESSION AX099776
VERSION AX099776.1 GI:13538810
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2109)
ANDERSSON, L., LOOF, C., KALM, F., MILAN, D., ROBC, A.,
ROGEL-GALLIARD, C., IANNUCELLI, N., GELLIN, J., LE ROY, P. and
CHARDON, P.
Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof
Patent: WO 0120003-A 3 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
ANDERSSON, Lelf (SE) ; LOOF, Christian (DE) ; KALM, Ernst (DE)
Location/Qualifiers
1..2109
/organism="Homo sapiens"
/db_xref="taxon:9606"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35799.1"
/db_xref="GI:13538811"
/translation="MRFOHTCTCDAMATSSKLVIFDTMLEIKKAFALVANGVRAAP
LMDSKQSFVGMLTITDFILVLRHYRSPLOVQIEIDQHKLETMRREIYLQCFPLVS
ISPDSLEAVVYTLKNIHRLPYLDPVSGVNLILTHKRLKFLHIGSLRPSFL
YRTIDLGITGTRDLAVVLEFAPILITADIDVDRVSLPYNVCGGVVGYXSEFDVI
HLAADOOTNHLDMVGEALRORTCLDEGLSCQPHESLGEVIDRIAREQVHRLVYDE
TOHLGVVSLSDIIOALVLSFAGIDALGA"

BASE COUNT 458 a 621 c 560 g 470 t
ORIGIN

Query Match 87.9%; Score 1447; DB 6; Length 2109;
Best Local Similarity 97.3%; Pred. No. 0;

Matches 1495; Conservative 0; Mismatches 35; Indels 7; Gaps 2;
OY 101 ttcttaagaagaagaacacgaagctcatgagcatcaacgaagctgtgaccagacgtcaaa 160
Db 1 TTCTTAGAGCAAGAAAACACAGCTCTATGGCATACACAGCTGTGACACAGCTAGAA 60
OY 161 agaatccgttggaagaagcgaagccttgaaatgagcaagagaggaagctggtggaag 220
Db 61 AGAATCCGTGGGAAGACGAGGCGCAAAAGCTTATAGATGACAAAGCCAGAAATGGTGA 120
OY 221 gaaagggaagcaccaggtcagggggaaggtcccggtlccagcgaagctgctgagtcacc 280
Db 121 GAAGGGGAGCCACACAGGTGAGGGGAAGGTCCCGGTCCAGGCCAAGCTGTGATGCAC 180
OY 281 gggctggaaggccaattccccaaagcacaaccccttgctcaagctgatacctgcgggtg 340
Db 181 GGGCTGAGGCGCAATTTCCCAAGACCACACCTTTGGCTCAACCTATCTTCCGGGGTG 240
OY 341 ggcactcaacacaaaggttgagactgctccctcctcagctgtacagctcagctgagc 400
Db 241 GGCATTCACACAAAGGAGGAGGAGTGCCTCCCTCTGACTGTACAGCTCAGCTGCAGCC 300
OY 401 tccaagcagagatgagtgtgagctggtccacggaagttcccaagcagaagccttggaagt 460
Db 301 TCCAGACAGATGATGTGAGCTGGCCAGAGTTCCAGCCACAGAGGCTTGGAGTGT 360
OY 461 gaagtagaaggcctgtctggaagaagaagcctgtgctctcctcccggaagcccat 520
Db 361 GAGCTAGAAAGCGCTGCTGGAAGAAGAGCCTGCCCTGTGCTGCTCCCGAGGCCCCATTT 420
OY 521 cccaagctgggtctggaatgacgaactgagaaacccggcaccagaatctacatgctc 580
Db 421 CCCAAGCTGGGCTGGGATGACGAACCTGCGGAACCCGGGCCAGATCTACATGGGCTTC 480
OY 581 atgcagcagacacacccgtcagcatgcatgagcaactagctccaagctgactcttgac 640
Db 481 ATGCAAGGACACACCTGCTACATGCAATGCAATGCAACTGCTCAACCTACTACTCTTGC 540
OY 641 accatgctggaagatcaagaagcctctctgtctcgtgtgagcaagcgtgtgcggaagc 700
Db 541 ACCATGCTGGAATCAAGAAAGCCTTCTTGTCTGTGTGCGCAAGGTGTGGGGAGCC 600
OY 701 cctctatggaagaagaagaagaagcctgtgtggaatgctgacacacacacacacac 760
Db 601 CCTATATGGAGCAGCAAGAAAGCAGACTTGTGGGATGCTGACCATCACTACTACTATC 660
OY 761 ctggtgctgcatcgctactacaggtcccccctggtccagatctatgagatgaaacaacat 820
Db 661 CGGTGCTGCATCGCTACTACAGGTCCCCCTGGTCCAGATCTATGAGATTTGAACACAT 720
OY 821 aagattgaagcctggaaggagaatctactctgaagcctgctcaagcctgtgtctcatc 880
Db 721 AAGATTGAACCTGGAGGAGATCTTACCTGCAAGCCTCTTCAACCCCTGTGCTCATC 780
OY 881 tctctaatgtagcctgttgaaagcctgttcaacccctcaagaagaacggatcatcgc 940
Db 781 TCTCTAATGATAGCTGTGTTGAAGCTGTCTACACCTTCATCAAGAACGGATCATGTC 840
OY 941 ctgcctgtcttgacccgggtgtcaggaacgctactcaacatctcaacacacacacacac 1000
Db 841 CTGCTGTGTTTGAACCCGGTGTCAAGCAAGCAAGTCCACATCTCACACAAACGCCCTG 900
OY 1001 ctcaagctcccgacaatcttggtctcctgctgcgcggccctcctcctacagcaact 1060
Db 901 CTCAAAGTTCGACATCTTGTGTTCCCTGTGCCCCGCCCTCTCTCTTACCGGACT 960
OY 1061 atccaagaattgggacatcgacatcccaagaccttggtgtcgtggagaagcagccc 1120
Db 961 ATCCAAGATTTGGGCAATCGGACATTCGAGACTTGGCTGTGGTCTGGAGACAGCACCC 1020
OY 1121 atctcactgactgagacatcttgtagaccggcgtgtgtctgactgactgctgtgtaac 1180
Db 1021 ATCCTGACTGCATGACATCTTGTGAGACCGGCGTGTGCTGACATGCGCTGTGTCAC 1080

Db	778	GTCTCCATCTCTCCCAATGACAGCCCTTTGGAACTCTCTAACGCCCTTCATCAAGACCGG	837
Qy	932	ATCCATCGCTGCCTGCTCTCTGAAACCCTGTCAGCAAGTACTCCATCCTCACAC	991
Db	838	ATCCAGCCGCTGCCTGGCTCTGGACCCCTGTCCTCCGGGCTGTGCTCCATCCTCACAT	897
Qy	992	AAAGCGCTGTCAAGTCTCCGTCGACATCTTGTCTCCCTGTCTGCCCCCTCCTCCTC	1051
Db	898	AAAGCGCTTCTCAAGTCTCTCTGACATCTTTTGGACACCTGTGCCCCCTCCTCCTC	957
Qy	1052	LACCGACATLCCAAAGATLTGGGATCGGACALCTCCAGAGCTGTGCTGTGTGTGAG	1111
Db	958	TACCGGACCATCCAAAGATTTGGGCATCGGACATATTCGACACTTGGCGGTGCTCGAA	1017
Qy	1112	ACAGCAACCATCTGTGCTGTGACCTGGGACATCTTGTGGACCGCGTGTCTGCACTGCT	1171
Db	1018	ACCGGCCCTCATCTGTGACCGGACGTGACATCTTGTGTGACCGGCGTGTGCTGCGCTC	1077
Qy	1172	GTGTCAAGCAATGTGTGCAAGTCTGCTGTGCGCTCATCTCCGCTTGAATGTATC	1231
Db	1078	GTGTGTCAAGCAAACTGTGACAGAGTATGTGGGCTCTACTCTGCTTATGTATGTATC	1137
Qy	1232	GTCTGCCAACAACCTACAGACCCCTGGGACATGATGTGTGGAGAGCCCTGTAGGCAAG	1291
Db	1138	GCTGCCAACAACAATACACACCTGTGACATGAAATGTGGAGAAACCCCTGAGCGAGCGG	1197
Qy	1292	ACACATLTGCTGTGAAGGAGTCTCTCTGACGAGCCCTGAGAGACCTTGGGGAGATCAT	1351
Db	1198	ACACTGTGCTGTGAAGCGCTCTTCTCTGACGCCCCACAGAGACCTTGGGGAGATCAT	1257
Qy	1352	GACCAAGATCTGCTCGGAGAGAGTACAGAGCTGTGTCTGTGTGTGTGTGTGTGTGTGT	1411
Db	1258	GACCGGATTTGCTCCGGGAACAGTGTGACCCCTGTGTCTGTGTGTGTGTGTGTGTGTGT	1317
Qy	1412	LTGGGCTGTGCTCTCTCTCTCGACATCTCTCTGAGCACTGTGTGTGTGTGTGTGTGTGT	1471
Db	1318	CTGGGCTGTGCTCTCTCTCTCTCTGACATCTCTGAGGCTGTGTGTGTGTGTGTGTGTGT	1377
Qy	1472	GTGTGCTGTGAGGCTGTGAGAGATCTGTGATCTCTCAATCTCCAGCAACCT-geacacctg	1530
Db	1378	GATGCTCTGTGGGCTGTGGAACCTTTGTCTTCAGGCGCCACTGTGACACACTGTGG	1437
Qy	1531	AAAGCAATGAAGGAGTGTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1590
Db	1438	AAAGCAATGAAGGAGGAGC-CTGTGAGCTGAGCTGTCAACTTCTCTCTCTCTCTCTCTCT	1496
Qy	1591	TCAGCTGTATTCAGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	1624
Db	1497	CTGTGCTCTTGTTCAGATGTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1530

RESULT	8				
LOCUS	AX099800		1873 bp	DNA	linear
DEFINITION	Sequence 27 from Patent WO0120003.				PAT 02-APR-2001
ACCESSION	AX099800				
VERSION	AX099800.1		GI:13538834		
KEYWORDS					
SOURCE					
ORGANISM					
	Sus scrofa				
	Euarchyotae; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.				
	1 (bases 1 to 1873)				
REFERENCE					
AUTHORS	Anderson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,				
	Regel-Galliard,C., Iannuccelli,N., Gellin,J., Le Roy,P. and				
	Charbon,P.				
TITLE					
	variants of the gamma chain of ampk, dna sequences encoding the				
	same, and uses thereof				
JOURNAL	Patent: WO 0120003-A 27 22-MAR-2001;				
	INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;				
	Andersson, Lelf (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)				
FEATURES					
source	1..1873				
	location/qualifiers				

```

/organism="Sus scrofa"
/db_xref="taxon:9823"
1..1395
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35800.1"
/db_xref="GI:13538835"
/translation="MSFLEGESRSPRAVYTTSSERHGDGQGNKASWHTNOEVEEED
GPGRREGQSPRAVSESTGOEATPPKAPLPLAQAAPLAVNDPNERRIILSPDCAADSA
DSNDHDLDLGIEFSASASGDELIVKPKAPCPSPFVLLPLGHDELQPCAGQVNYH
HMDHEHTCYDAMATSSKIVPDTMLKIEKAPFALYANVRAPRLMDSKQSFQGLT
TDFLLIYHRYRSPLOYVEIIEKHIEYWRILYLDGCKPLVLSIPNDSLEEAYVAL
KNRIHRLPVLDPYSGAVLHLIIFHRFLKFLHI FGLLDRPSFLRTIODGLIGFRDIL
AVVLETRALITLALDIPDRVSALPLVNETQVVGFSRFPV IHLAAQOYTNHLMDSN
GEALVQRITLLEGVLSQCPHEPTLGEVYIDRIIVREOVHRLVLDVETQHLGVVSLSDILG
ALVSPALDILG"
BASE COUNT      382 a      580 c      535 g      376 t
ORIGIN

```

Query Match	69.28;	Score 1140.4;	DB 6;	Length 1873;
Best Local Similarity	85.7%;	Pred. No. 3.3e-250;		
Matches 1315;	Conservative	0;	Mismatches 211;	Indels 8;
				Gaps 4;

Qy	atgaagcttcttaagaaagaanaaacagcaatcatgtgcaatccacagctgtgacagcagc	154
Db	1 attagcttcttaagaaagaagagagacccttattatggccattccggagcttaccacacg	60
Qy	155 lcaagaagaatccgttggaacaagcagcgaagccctltagaatggaagcagaagtcg	214
Db	61 tcgaaagaagccatgtagggaccaggggaacaaagccctttagatggacaaagcagagcat	120
Qy	215 gtggaaagaaggagagccaccaggtcagggggaaggtlccccgtlccagcagctgtgag	274
Db	121 gtagaagaagggggggcctccggcccgaaaggaaagtcctccactccagccagcatgttgag	180
Qy	275 tccacagggctggaagcgcacatcccccaagacacaaccccttggcccaagctgatac---t	331
Db	181 tccacggggcagagagccacatttccccaagccacaccccttggcccaagccgctcttg	240
Qy	332 ggcggagggtgggaatccaccaaagaggttggagctgcttcccctctgactgtacagctca	391
Db	241 gccgaagtgagacaaaccccccaacaaagggggacatcttcccttgcagctggacgctca	300
Qy	392 gctgcagcgtccagacacagatgatgtggagctgcgcagagagtlcccaagcacaagaagcc	451
Db	301 gcctcgacttccaaacacagaccatttgatcttgggcattagagtttccagccttcggggcg	360
Qy	452 tgggagtgtagctcagaagcgtcgtctlgaagaagagcctgtgctctgtgcttccgcag	511
Db	361 tggggggatgagact--tgggctgtgtggaaggaagccagccgctggtccatccccagag	417
Qy	512 gccccattcccaagcgtgagctgtggaatgaagaaactgaggaaccggagcccaatcac	571
Db	418 gtagctgattaccagagcctgggctggagatattagctgcgaamaagccggggccacagcttac	477
Qy	572 atgagctcacaagcagaagacaaactgtctcagatgcgaatgagactagctccaagctagc	631
Db	478 atggactttatgacagagacacacctgcgtacatgcattggcgagcacactccaaactgtgc	537
Qy	632 atcttcgacacacatctgtgagatcaagaagagccttcttgcctgtggtgccaaggtg	691
Db	538 atcttcgacacacatgctgagatcaagaagagccttcttgcctgtggtgccaagaggtg	597
Qy	692 cggagagccctctatggagcagcaagaagcagagcttgtlbggaatgtgacatcaact	751
Db	598 cgagggagcacttggtagcagcagcaagaagcagagcttgcggggagatgtagacatcaca	657
Qy	752 gacttcaccccggtgctgacatgcgtactacaaggtccccccctgtgccaagatctatgaatt	811
Db	658 gacttcaccccggtgctgacatgcgtactacaaggtccccccctgtgccaagatctatgaatt	717
Qy	812 gaacacataaagatttgagaccttgaggaggaatcactacctcaagctgtgcttcaagcccttg	871

|||||
Db 718 GAAAGAACATTAAGATGACACCTGGAGGAGATCTACCTTCAAGGCTGCTCAAGCCTCTG 777
Oy 872 gtcccatctctcctaatgataagcctgtcttgaaagctgtctaacacctcaataagaccg 931
Db 778 GTCTCATCTCTCCCAATGACAGCTGTGGAAGCTGTCTACGCCCTCATCAAGAACGG 837
Oy 932 atccatccctgagcctgtctcttgagaccggtgtcgaagcaagctactcaatcccaaac 991
Db 838 ATCCACCCCTCTCCGCTCTGACCTCTCTCCGGGCTGTGCTCCACATCTCTCACACAT 897
Oy 992 aaacgcctgctcaagctcctgacacatctctgtgacccgctgtgctgacccctctctc 1051
Db 898 AAGCGGCTTCTCAACTTCTCTGACATCTTTGGACACCTGCTGCCGCCCTCTCTCTC 957
Oy 1052 taacgactatccaaagtattgagcctgagacatccgaagactgtgctgtgctgag 1111
Db 958 TACCCACCATTCAGAAATTTGGGATCGGCACATTCGAGACTTGGCGCTGTGGAA 1017
Oy 1112 acaagaccatccctgactgacatctctgtgacccgctgtgctgacactgct 1171
Db 1018 ACGCGCCCATCTCTGACCGCACTGACATCTTGTGACCGGCTGTGCTGCTGCT 1077
Oy 1172 gtgtcaagaagatgtgtcaggtcgtgagccctatctccgctgtgactgactgactg 1231
Db 1078 GTGTCAACGAAGACTGACAGCTAGTGGGCTCTACTCTGCTTGTGATGATCAGCTG 1137
Oy 1232 gctgcccgaagaacctataacacactgagacatgtgtgtgagagaagccctgaagcaga 1291
Db 1138 CTTGCCCAACAAATACACACCTTGAGATGATGAGAAAGCCCTGAGGACGCG 1197
Oy 1292 acaactatgtcgtgagagagctctctcgcagaccgcccagagagctgtgaggaagtgc 1351
Db 1198 ACAGTGTCTGTGAAGGCGCTCTTCTGCGACGCCACGAGACTTGGGGAACTCATTT 1257
Oy 1352 gacagagatgtctcgtgagacagtlacagagctgtgtgactgtgagcagagaccagatctc 1411
Db 1258 GACCGGATGTTCGGGAGACAGTGCACCGCTGTGCTGTGATGAGAACCCACACACTT 1317
Oy 1412 ttgagcgtgtctcctctcctcctcgcagacatctctgagcactgtgtgtcagagccgtcgtgcata 1471
Db 1318 CTGGCGCGTGTCTCTCTCTCTGACATCTTCAAGCTGTGTGCTCAAGCTGTGGAATTT 1377
Oy 1472 gatgcctcgtgagccttgagaaagatctgagctcactcaatcccaagcactc-gcaacactg 1530
Db 1378 GATGCGCTCGCGGCGCTGAGAACCTTGGAACTTGTCTGAGCGACCTGGCACACCTCG 1437
Oy 1531 aagccaatgaaggaagcactgagaaactcaagctcactcctcccccaccccaattgtcgtgt 1590
Db 1438 AAGCCAAGTGAAGGAGC-CGTGAGACTCAGCTCTCACTTCCCTCAGGCCACTTGTGCTGT 1496
Oy 1591 tcaagctatgactcagtagagctcgtccctgagcc 1624
Db 1497 CTGGCTCTTTCAGTGAAGCTCCGCCCGGGGCC 1530

RESULT 9
AX099774 1867 bp DNA linear PAT 02-APR-2001
LOCUS Sequence 1 from Patent W00120003.
DEFINITION AX099774
VERSION AX099774.1 GI:13538808
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 1867)
Andersson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Rogel-Gaillard, C., Iannucci, J. N., Geilvin, J., Le Roy, P., and
Chardon, P.
Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof

JOURNAL Patent: WO 0120003-A 1 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES
source
1. 1867
/organism="Sus scrofa"
/db_xref="taxon:9823"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="GAC35798.1"
/translation="MHFMDHTCYDAMATSSKLVIFDTMLEIKAFPALVANGVRAP
LMDSKOSFVGLTIDFDILVLRVRYRSPVLOVYEIEHKIETWREYLOGCFPLVS
ISPNDLSFEAVVALIKIRIHRPLVDPSVSAVLIHKLKFLHIFGLRPSFL
YRTIODIGTGFRLAVLETAIILFALDIFVDRVSALEPVNETGVVGLYSRFDVI
HLAAOQTNHLDMNVGBALRPTLCEGLVISCQPHFTLGEVIDRIYRBOVHRLVLVDE
TQHLGVSLSDILQALVLSRAPDILGA"
BASE COUNT 380 a 583 c 529 g 375 t
ORIGIN

Query Match 68.9%; Score 1134.4; DB 6; Length 1867;
Best Local Similarity 85.7%; Pred. No. 7.7e-249;
Matches 1309; Conservative 0; Mismatches 211; Indels 8; Gaps 4;

Oy 101 ttcctagagcaagaacaagcagctcatalgaccatccagcctgtgacagcagctcagaa 160
Db 1 TTCTTAGAGCAAGAGAGAGAGCCGTTTCATGGCATCCGCACTGTACACACAGCTCAGAA 60
Oy 161 agaatccgtggaagaaaggaagggcccaagccttgagatgagcaaggaagcgtgtgagag 220
Db 61 AGAAGCCATGGGAGCCAGGAGAACAGGCCCTTAGATGACAAAGGAGAGATGTAGAG 120
Oy 221 gaagggagccacaaggtcagaggtcccggtccagggccagctgtgagtcacc 280
Db 121 GAAAGGGGCGCTCCGGGCGGAGGAAGTCCACAGTCCAGGCCAGTGTGAGTCCACC 180
Oy 281 ggcgtgagagccacatctccccaagacacacacctgtgctcaagctgac---tgccggg 337
Db 181 GGGCAGGAGGCCACATTTCCCAAGGCCACACCTTGGCCCAACCGCTCCCTTGGCGGAG 240
Oy 338 gtggcactccaccaagaaggtggaactgctccctctgacactgtacagctcagctgca 397
Db 241 GTGAGACAAACCCCAACAGAGCGGAGACATCTCCCTCTACTGTGACGCTTACGCTCC 300
Oy 398 ggtccagcagatgattgtgagactgtgccaagagttcccaagcacaagagcctggag 457
Db 301 GACTCCAAACACAGACATCTGATCTGGCATGAGAGTTTCACACCTCGCGGCGTGGGG 360
Oy 458 tgtgagctagaagagcctgtcgtgagaagagagcgtccctgtgctgtcccgcaagccca 517
Db 361 GATGAGCT--TGCGCTGTGTGAAGAAAGAACGCCCGCCGTCACATCCAGAGGTGTGTG 417
Oy 518 ttcccaagctgtggtgagtgagaaactgtcggaacccggtgcccagatctacatgctgc 577
Db 418 TTACCCAGGCTGGGCTGTGGATGTGATGACTGACAAAGCCGGGGCCCAAGGCTTACATGAC 477
Oy 578 ttcattcaagagacacactgtactgatagtcattgtgcaactagcttccaagtagtcatctc 637
Db 478 TTTCATCAGAGACACACTGCTACGATGCCATGGACAGCTCAAACTGGTCATCTTC 537
Oy 638 gacacatgctgtagaatacaagaagagcctctgtgctcgtgtgagcaagagtggtcgagca 697
Db 538 GACACATGCTGTGAGATCAAGAGGCTTTTGTGCTGTGGCTGTGGCCAAAGCGCTCCGAGG 597
Oy 698 gccctctatgtgagcagaagaagcagagctgtgtgagatctgtgacatactgactc 757
Db 598 GCACCTTTGTGGGAGACGAAGAAGAGAGACTTCTGTGGGATGCTGACATTCACAGACTTC 657
Oy 758 atctgtgtctgcatcgtactacaggtcccccctgtgtcagatctatgtagatgtgaaca 817
Db 658 ATCTTGCTGTGACCGCTATTACAGGTCCCGCTGTGTCAAGATGTACGAGATGTGAAGA 717

QY 818 calaagattgagacttggaaggaatctacctgcaagctgctcaagcctctgtctcc 877
 |||||
 Db 718 CATAGATTGACACCTCGAGGAGATCTACCTTAAGCTCTTCAAGCCTCTGCTGCC 777
 QY 878 atctctcaatgtaagactgcttgaaagctgtacacacccatcaagaacccagatcat 937
 |||||
 Db 778 ATCTCTCCCATGACAGCCTTTTGAAAGCTTCTACGCCCTCATCAAGACGGATCCAC 837
 QY 938 cgcctgctgcttcttgaaacccggtgctgaagcaagctactccacatctcacacacacgc 997
 |||||
 Db 838 CGCCTCGCGCTCTGAGACCTGTCTCCGGGCTGTGCTCCACATCTTCACATCAAGCGG 897
 QY 998 ctgctcaagttctcgcacatcttctgtctccctgtgctcccgccctctctctctacgc 1057
 |||||
 Db 898 CTCTCTCAAGTCTCTGACATCTTTGGACACCTGTGCTGCCGCCCTCTCTCTTACCGC 957
 QY 1058 actatcaagaalttgagcaatcgcacatctcgaagacttgctgtgctgctgaagacaga 1117
 |||||
 Db 958 ACATCCAGATTGGGCATCGGACATTCGACATTCGGCCGTGGTGGTGGAAACGGCG 1017
 QY 1118 cccatctgactgactcgtgagacatcttcttgagccgctgtgctgactgctgtgctc 1177
 |||||
 Db 1018 CCCATCTTGACCGCAGCATCTTGTGACCGCGGTGTGTCTGCGCTGCTGTGCTG 1077
 QY 1178 aacgaatgltgcaagctgctgagccctatctccgccttgatgatacactgctgctgctc 1237
 |||||
 Db 1078 AACCAAACTGACAGTAGTGGCTCTTACTCTGCTTTGATGATCAACCTGGCTGCC 1137
 QY 1238 caagcaactcaac 1297
 |||||
 Db 1138 CAACAAACATACACACACCTGAGCATGATGTGGAGAACCCCTGAGAGCAGCGACACTG 1197
 QY 1298 tctctggaaggaagctctctctctccgagcccaagagaagcttgaggagaatgacagag 1357
 |||||
 Db 1198 TGTCTGGAAGCGCTCTCTTCTGTCACCCACACACCTTGGGGAAGTCAATGACCGG 1257
 QY 1358 atgctcgtggaagcagctacacacacacacacacacacacacacacacacacacac 1417
 |||||
 Db 1258 ATTGTCCGGGAGACAGTGCACCCCTGTGCTGTGATGATGACCCGACCTTCTGAGC 1317
 QY 1418 gtggctctccctctcgcacatctcctcaagcagctgctgctcagcctgctgacatgac 1477
 |||||
 Db 1318 GTGTGTCTCTCTCTGACATCTTCAAGCTGTGTGCTGACGCTGCTGTAATGATGCC 1377
 QY 1478 ctgagggccggaagaagctcgtcgaatctccatcccaagcagcct- gaacacttgagac 1536
 |||||
 Db 1378 CTGCGGGCTGAGAACCTTGGAACTTGTCTCTGAGGCCACCTGGACACCTTGGAAACCA 1437
 QY 1537 atgaaggaagctgagaactcaagctcatalctctcccccacccatctgctgctcagct 1596
 |||||
 Db 1438 GTGAAGGAGAC- CGTGACATCTACCTTCCCTTCACACCCGACCTTGTGTGTGTGTGT 1496
 QY 1597 atgattcagttgaagctctgctccctgggccc 1624
 |||||
 Db 1497 CTTGTTCAAGTGAAGCTTCCGCCCGGGGCC 1524

RESULT 10
 RNAMPKAM 1328 bp mRNA linear ROD 25-FEB-1998
 LOCUS R. norvegicus mRNA for gamma subunit of AMP-activated protein
 DEFINITION kinase.
 ACCESSION X95578.1 GI:1185270
 VERSION X95578
 KEYWORDS AMP-activated protein kinase; gamma subunit.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1328)
 AUTHORS Woods, A., Cheung, P. C., Smith, F. C., Davison, M. D., Scott, J.,

TITLE Bert, R. K. and Carling, D.
 JOURNAL Characterization of AMP-activated protein kinase beta and gamma
 MEDLINE Subunits. Assembly of the heterotrimeric complex in vitro
 REFERENCE J. Biol. Chem. 271 (17), 10282-10290 (1996)
 96215327
 AUTHORS 2 (bases 1 to 1328)
 TITLE Carling, D.
 JOURNAL Direct Submmission
 Submitted (07-FEB-1996) D. Carling, MRC Clinical Sciences Centre,
 Department of Molecular Medicine, RPMs, Hammersmith Hospital,
 DuCane Road, London, W12 0NN, UK

FEATURES
 source
 1..1328
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 37..1029
 /function="binds to alpha and beta subunits of AMPK,
 forming a heterotrimer"
 /codon_start=1
 /product="AMP-activated protein kinase gamma"
 /protein_id="CAA64831.1"
 /db_xref="GI:1185271"
 /db_xref="SWISS-PROT:P80385"
 /translation="MESVAEESAPAPENHSETPESNSVYTTTMRKSHRCYDLPIS
 SKLVFEDISLQKKAFFALVTVGNRAAPLMDSKOSFVGMLTITTOFINLHRYRSAL
 VOILELEHRIETMRBRYLQDSFKPLVCLSPNVAFLDVASSLIRMKIRTPYIDPESG
 NPLVILTHKRIKFLKLFTEPRKPEFNSKSLSEELQIGYANIAMVARTTTPYVNLGI
 FVQHRVSLPVDDEKRVVDYISKFDVILAEKFTNNLDVETVALOHRSHYFEGVL
 KYLHEETLEAILNRLVEAEVHLVYVDEIDVVKGIYSLSLIDALVLTGEEKP"

BASE COUNT 330 a 337 c 340 g 321 t
 ORIGIN

Query Match 25.2%; Score 414.8; DB 10; Length 1328;
 Best Local Similarity 64.9%; Pred. No. 2, 3e-84;
 Matches 614; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 512 gccccatctcccaagctggtggtgagtgagcaactcgcggaacccgcgcccagatctac 571
 |||||
 Db 61 GCTCCAGCTCCGGAGATGAACACTCTCAAGACCCCGGAAATCGAAGATGATGTGTC 120
 QY 572 atgagcttaatgagaagacacactgctcagatgcagatcgaactcgaagctgac 631
 |||||
 Db 121 ACCACCTTATGAAGTCTCTCTGCTGCTGATGACCTGATCCCAAGCTCAAGCTGGTG 180
 QY 632 atcttcgaacacatctgtgagatcaagaagcctctctgtctgtgtggtgcacagtgctg 691
 |||||
 Db 181 GATTGTGATACCTGCTGCGAGGTGAAGAAAGCCTTCTTGCCCTGTGACTAAGCGTGT 240
 QY 692 cggagagccctctatagggacagcaagaagcctctgtgtggagtgctgacatcact 751
 |||||
 Db 241 CTTGCTGCCCTTTGTGGGATGATGAAGACAGACTTTGTGGGCACTCTGACATCACT 300
 QY 752 gactcatcctgtgtgtgctgactcactacagctcccccctgctgcaagatcatalagatt 811
 |||||
 Db 301 GATTATATCAATATTCTGACACCGATACTCAAGTCAAGCCTGTGACAGATCTATGACTG 360
 QY 812 gaacaacataagattgagaactcgtgaggaagatctactcctgcaagctgtctcaagcctctg 871
 |||||
 Db 361 GAGGAGACACAGATAGAGACTTGGAGAGAGCTTCACTCTCAAGACTCTTTAAGCACATT 420
 QY 872 gctcactctctccatagatagctgtgttgaaagctgtcttaacacctcaagaacgg 931
 |||||
 Db 421 GTCTGATTTTCTCAAAATCTCCAGCTTGTGATGCTCTTTCATTAATTGGAATAAG 480
 QY 932 atccatcgctgctgctgtcttgaccggtgtcaggcaagctactcaactcctcacac 991
 |||||
 Db 481 ATCCACAGGCTTCCAGTTATTACCCGGAGTCAAGCAACCTTGTATCTTTTACTAC 540
 QY 992 aaagcctgctcaagttcctgacacatcttgctccctgctgctccggccctctctc 1051
 |||||
 Db 541 AAGCGATCTCAAGTCTCTCAAGTTGTTATCACTGAGATTCCCAAGCCGAATTCATG 600

Oy	1052	taccgacatcccaagatttgggcatacggcacatccgagacttgctgtgtgctggaag	1111
Db	601	TCCTAAGTCTCTGCAAGAGACTACAGATGTGGCACTACGCCAAATATTGCCATGGTCCGTACC	660
Oy	1112	acagaccccacttcctgacctgacctgtagacaccttttgtaggaccggcgctgtgctgcatgct	1171
Db	661	ACTACACCTCGTCTATGTGTGGCTTGCGGCATCTTTGTATACAGCACCGAGTCTCGCCTTGCTT	720
Oy	117	gtcgtacaacqaatlgtgtltaagtcgtgtgggccccctclatcccgctttgatgtgtaacctg	1231
Db	721	GTCGTGATGAGAAMAAGGCCGTGTGGTGAGCATCTACTCCAACTTGATGTATTAATTTG	780
Oy	1232	gctgcccgagaacctatacagaccactgtgacatgtgtgtggagaagcccttgagcagaag	1291
Db	781	GCAGCAGAAAACAACATAACAACACTTAGTGTGTGTACAAAAAGCCCCCTACAGCACCGG	840
Oy	1292	aacacatgctgtgaaggagctctcttccttcggccagccccagagaagcttggggaaagtatc	1351
Db	841	TCACACTACTTTCGAGAGGCTGTTCTCAAGTGTCTACCTACATGAACATCTGMAAACCAATCATC	900
Oy	1352	gacacgattgtctcgggagcagglatacagagcctcgtgtcagltgtgagaagaccagatctc	1411
Db	901	AATAGACTGTGCAAGACAGAGCTTCACCGCTGTGGTGCTGTGATGAACATGACCTGGTGC	960
Oy	1412	tggggcggtgtctccctccctcgcacatcccttcaggcactgtgtctca	1457
Db	961	AAGGCAATTGTATCGTCTGTGACATCTTACAGGCTGTGTCCTCA	1006
RESULT_11	RNU42413	1550 bp	mRNA linear ROD 30-MAY-1996
LOCUS DEFINITION	Rattus norvegicus 5'-AMP-activated protein kinase, gamma-1 subunit mRNA, complete cds.		
ACCESSION VERSION	U42413		
KEYWORDS	U42413.1 GI:1335859		
SOURCE ORGANISM	Norway rat. Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Rattus.		
REFERENCE	1 (bases 1 to 1550)		
AUTHORS	Gao,G., Fernandez,C.S., Stapleton,D., Auster,A.S., Widmer,J., Dyck,U.R., Kemp,B.E. and Witters,L.A.		
TITLE	Non-catalytic beta- and gamma-subunit isoforms of the 5'-AMP-activated protein kinase		
JOURNAL MEDLINE REFERENCE	J. Biol. Chem. 271 (15), 8675-8681 (1996) 96224074		
AUTHORS	2 (bases 1 to 1550) Gao,G., Widmer,J., Auster,A., Stapleton,D.S., Kemp,B.E. and Witters,L.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-DIC-1995) Lee A. Witters, Medicine/Biochemistry,		
MEDLINE REFERENCE	Dartmouth Medical School, N. College St., Hanover, NH 03755-3833, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..1550		
CDS	/organism="Rattus norvegicus" /db_xref="taxon:10116" <1..972 /codon_start=1 /product="5'-AMP-activated protein kinase, gamma-1 subunit" /protein_id="AAC52580.1" /translation="SAPAPENSHOETPESNSSVYTFEMKSHRCYDLPITSSKLIVPDTSLOVKRAFPALVTNGVRAPLMDMSKOSPQMCTIDPFILNIIRAKYSKLYOTIELEEHKITEREVLIODSPKPLVCISGNALSIPAVSSLINRKIHLRPVIDREGNLTILT HKRIKLRLKLFTEFRPKRPBPMKSLSLELDIGTTANIAMKRTPPVVALDIPIQHRVAS ALPVEDKGRNVDIYSKFEDIYNLAETENTNNLVDSYKVALQSDHRSHEFGLCYLHET LEAIINLVAEVHRIWVDHDHNQIVSLISDLQALVTLTGSEKKR"		

BASE COUNT	371 a	398 c	394 g	387 t	
ORIGIN					
Query Match	25.2%	Score 414.8:	DB 10:	Length 1550:	
Best Local Similarity	64.9%	Pred. No. 2, 3e-84:			
Matches 614:	Conservative	0:	Mismatches 332:	Indels 0:	Gaps 0:
QY	512	gccccaattcccaagctcggcgtcggatgaaacgaacatcgcggaacccgcgcagacatctac	571		
DB	4	GCCTCAGCTCCGGAGATGAAACACTCTCAAGAGACCCCGAAATCGAATCAGTATGTTGTAC	63		
QY	572	atgcggttatcgaaggagacaacatctcgtcgtatgcattcgaatgcacatagctacgcacgtc	631		
DB	64	ACCACCTTATGATGAGCTCTATCGCTGCTATGACCTGATCCCAACAAGCTCCAAAGCTGGTG	123		
QY	632	atcttcgacacacatgcctgcgtagalcaaaagagcctcttcctcgtcgtcgtcgaagcgtg	691		
DB	124	GATATTGATACCTTCGCTGCGAGGTAAAGAAAGCTCTTTGCCCTGGTGACTAACGTTGT	183		
QY	692	cgggcagccctctcatalyggagcaagaaagcagaagctcttgaggatgctgaacacact	751		
DB	184	CGTGCCTCCCTTTGTGGATAGTAAAGAGCAGAGCTTTGTGGGCTGCTGACCATCACT	243		
QY	752	gacttactcgtgctgtagatcgtcgaactaaagctccccccttgctcagaactatagatt	811		
DB	244	GACTTATATTAATTTTGTGACCGGATCTCAAGTCAAGCCCTGTGCGAGATCTATGAATCG	303		
QY	812	gaacaacataaagatltgagaccctggagggagagatcactcgaagctcgaagcctcgtc	871		
DB	304	GAGGACACACAGATAGAGACTTGGAGAGAGGCTACCTGACCAAGACTCCTTTAAGCCACTT	363		
QY	872	gtctccatctctcctaaatgataagctglttgaagctcgtctacacccctcaagaacg	931		
DB	364	GTCTGATTTCCTCAAAATCCAGCTTGTTCGATGCTCTCTTCAATTAATTCGAAATAG	423		
QY	932	atccacgcctcgcgtcttgaaccggtgcgaagcaagtaactcaactcctacacac	991		
DB	424	ATCCACAGGCTTCAGATTATGACCCGGAGTACGGCMAACCTTGATACATTTTACTCAC	483		
QY	992	aaagcgcgtcgaagctcctgcacacatlttgltccctcgtcgcgcgcgcctcctctc	1051		
DB	484	AAGCGGATCTCTCAAGTTCCTCAAGTGTGTTATACATGAGATTCCTCCCAAGCCGAAATTCATG	543		
QY	1052	taccgacatcccaagatlttgagcaltcggacatccgaactccgaacttgctggtgctgag	1111		
DB	544	TCTAAGTCTCTGGAAGAGCTTACAGATTGGCACCTACGCCAATATTGCCATGGTCCGTACC	603		
QY	1112	acagaccccatccttgaactcgaactgacatctcttgtagaccggtcgtctcgaactgctc	1171		
DB	604	ACTACACCTGCTATATGTGGCTCTGGGCACTTTGTATACAGACACGAGTCCGCCCTTGGCT	663		
QY	1172	gtgtgcacaaagatgtgtgtcaagtgctgtgggcctctatccgcctctgtagtgaactcgt	1231		
DB	664	GTGTGTGATGAAAGGGGCTGTGGTGCACTACTCTCCAAATTTTGATGTGATTAATTGG	723		
QY	1232	gtctgcacaaacaaactcaaacacactcgtgacatgagtgctggagaaagcctcgaagcagag	1291		
DB	724	GCAGCGAAGAAACATACATACACAACCTAGATGTGTGTGTGACAAAGCCCTACAGCACGG	783		
QY	1292	acaactatgtctggaggaagctcttcctcgtcgcgaagcccaagagagctctgggggaagtgc	1351		
DB	784	TCACACTACTTCGAGGGGTGTTCTCAAGTGCTACCTACATGTGAGACTCTGAGAACATATC	843		
QY	1352	gacaaagatgtcctcggagcaagatlaaacagcgtgtgtcgaagtggagagaccagacatctc	1411		
DB	844	AATTAACGTGTGGAAGACAGAGGTTACCGCTGTGTGTGTGTGTGATGAACATGACGTGTC	903		
QY	1412	ttgggcgtgtgtctcctctccgacatccctcaaggaactcgaagcgtgtgctca	1457		
DB	904	AAGGCATTTGTATCGCTGTGTGACATCTTAAAGGCTGTGGTGCTCA	949		

RESULT 12
AX281579 989 bp DNA linear PAT 02-NOV-2001
LOCUS AX281579 Sequence 2 from Patent WO0177305.
DEFINITION AX281579
ACCESSION AX281579
VERSION AX281579.1 GI:16608830
KEYWORDS
SOURCE human.
OKANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Anderson, L., Luthman, H. and Martlund, S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 2 18-Oct-2001;
Arexis AB (SE)
FEATURES
source location/Qualifiers
1..989
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 229 a 306 c 286 g 168 t
ORIGIN
Query Match 24.7%; Score 407.2; DB 6; Length 989;
Best Local Similarity 99.3%; Pred. No. 1,3e-82;
Matches 409; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 247 aggtcccggtccagcagcagctgtcagtcacacgggtgtgagggcacatccccaagac 306
|||||
DB 540 AGGTCCCGGTCACAGGCGAGCTGTAGTCCACCGGCTGGAGGCCACATTTCCCAAGAC 599
OY 307 caccacctgtcagctagctgtcctcgggggtgggacctcaccacaagggtgtgagctg 366
|||||
DB 600 CACACCTTGGCTCAGCTGATCTGCGGGGTGGGCACTCCACCACAGGCTGGGACTG 659
OY 367 cctcccccctgactgtacagcctcagctcagcagcagcagcagatgtgtggaagctgc 426
|||||
DB 660 CCTCCCTGTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGAGACTGCG 719
OY 427 caccgagttcccaagcacaagagggctgtgagctagaagctcgtcgtggaagag 486
DB 720 CACGGAGTTCCAGCAGCAGAGGCGCTGGAGCTGTAGCTAGAGGCCCTGCTGGAAGAGG 779
OY 487 gcttcgctgtcgtcgtcctcgcagagcccatlcccaagctgtgagtgacgaact 546
|||||
DB 780 GCCTGCCCTGTGCGCTGTCCCGCAGGCCCAATTTCCAAGCTGGCTGGATGACGAACCT 839
OY 547 gcggaacccggcgccagatctacatgctcaltcaltgcaagagcacacctgtacgaltgc 606
DB 840 GCGCAACCCGCGCCCGCCAGATCTACATCGCTTCATCCAGAGCAGACCTGCTAGATGC 899
OY 607 catggcaactgtcctcgaagctagatcattcttgacacccatgtgtgagatcaag 658
DB 900 CATGGCAACTGTAGCTCCAGTAGTATCTTGCACACCATGTGGAGGTGAGG 951

RESULT 13
AC073128 196554 bp DNA linear HTG 21-FEB-2001
LOCUS AC073128/c Homo sapiens chromosome 2 clone RP11-64705, WORKING DRAFT SEQUENCE,
DEFINITION 17 unordered pieces.
ACCESSION AC073128
VERSION AC073128.3 GI:13027579
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
OKANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Waterston, R. H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished

REFERENCE
AUTHORS 2 (bases 1 to 196554)
TITLE Waterston, R. H.
JOURNAL Direct Submission
Submitted (08-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 21, 2001 this sequence version replaced g1:8469048.
COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Project Information
Center project name: H_NH0647005
Summary Statistics
Sequencing vector: p13: 98%
Sequencing vector: plasmid: 0%
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187795 bases at least Q40
Consensus quality: 190513 bases at least Q30
Consensus quality: 192099 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 194954; sum-of-contigs
Quality coverage: 5.58 in Q20 bases; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1157: contig of 1157 bp in length
* 1158 1257: gap of unknown length
* 1258 3600: contig of 2343 bp in length
* 3601 3700: gap of unknown length
* 3701 5103: contig of 1403 bp in length
* 5104 5203: gap of unknown length
* 5204 8524: contig of 3321 bp in length
* 8525 8625: gap of unknown length
* 8625 11855: contig of 3232 bp in length
* 11857 11956: gap of unknown length
* 11957 15783: contig of 3827 bp in length
* 15784 15883: gap of unknown length
* 15883 21906: contig of 6023 bp in length
* 21907 22006: gap of unknown length
* 22007 28887: contig of 6881 bp in length
* 28888 28987: gap of unknown length
* 28988 35255: contig of 6268 bp in length
* 35255 35355: gap of unknown length
* 35355 44642: contig of 9287 bp in length
* 44643 44742: gap of unknown length
* 44743 58275: contig of 13533 bp in length
* 58276 58375: gap of unknown length
* 58376 73816: contig of 15441 bp in length
* 73817 73916: gap of unknown length
* 73917 92140: contig of 18224 bp in length
* 92141 92240: gap of unknown length
* 92241 113337: contig of 21097 bp in length
* 113338 113457: gap of unknown length
* 113458 130325: contig of 16888 bp in length
* 130326 130425: gap of unknown length
* 130426 149287: contig of 18662 bp in length
* 149288 149388: gap of unknown length
* 149388 196554: contig of 47167 bp in length.
Location/Qualifiers
1..196554
/organism="Homo sapiens"
/db_xref="taxon:9606"

[illegible]

	Query Match	24.7%	Score 407.2;	DB 2;	Length 196554;
	Best Local Similarity	99.3%;	Prism No. 7.8e-83;		
	Matches 409;	Conservative 0;	Mismatch 83;	Indels 0;	Gaps 0;
OY	247	aggctcccggtgcccaagccagcagctgctgagtcagtcacccggcgctgtagggccacatccccaagac	306		
Db	62887	AGGTCCCGGTCAGGCCAGCGCTGTAAGTCACCCGGGTGAGGCCACATTCCCAAGAC	62828		
OY	307	caaacctcttgctcaagctgatactctgcgggggttgtagcactccacaacaaggctgtagctg	366		
Db	62837	CACACCTTTGGGCTCAAGCTGATCTGCCGGGTGGGCACTCCACCACAAGGTGGGAGCTG	62768		
OY	367	cctcccccctgactgttaagcctcagctcgcaagggcccaagcaagttgattgtagaactgac	426		
Db	62767	CTTCCCTCTTCACTTTACAGCTTCAGCTGCACAGCTCCACACAGATGATGTGAACTGGC	62708		
OY	427	cacgagatctcccaagccacaagaagcctctggagctgtagctlaagaagacccgtctgaagaag	486		
Db	62707	CACGAGATTCCACGACACAGACGCTGGGAGTGTACGTAGAAAGGCTGCTGGAAGAGG	62648		
OY	487	gctgacctgtgtgctgtgtcccgcaagggcccaatctcccaagctgggctgtgtagtaagcaact	546		
Db	62647	GCCTCCCTGTGCTCTGTCTCCCGACGCCCCATTCTCCAAAGCTGGGCTGGGAAGAGAACT	62588		
OY	547	ggggaacaacccggcgcccaagatctatagtcgctcttcaagagagacaacctgttaagatgc	606		
Db	62587	GGGGAACCCGGCGCCCAAGATTACATGCGCTTTCATGCAAGGACACCACTGCTTACGATGC	62528		
OY	607	catggcaactatgctccaagctagtcactcttgacacacatgctggagatcaa	658		

Db 62527 CATGGCACTAGCTCCAGCTAGTATCTTCGACACCATGTGGAGGTGAGG 62476

RESULT	14	1623 bp	RNA	linear	ROD 30-MAY-2001
LOCUS	AF036535				
DEFINITION	Mus musculus AMP activated protein kinase mRNA, complete cds.				
ACCESSION	AF036535				
VERSION	AF036535.1	GI:2766684			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 1623)	Shamsadin, R., Jantsan, K., Adham, I. and Engel, W.	Cloning, organisation, chromosomal localization and expression analysis of the mouse <i>Prkag1</i> gene	Cytogenet. Cell Genet. 92 (1-2), 134-138 (2001)

REFERENCE 2 (bases 1 to 1623)

TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) AG.Engel, Humangenetik, Gosslerstr.12 d,

FEATURES	Location/Qualifiers
source	1. .1623

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_type="testes"
43. .1035
CDS

```

BASE COUNT	383 a	414 c	416 g	400 t	10 others
ORIGIN					

Query Match 24.1%; Score 396.8; DB 10; Length 1623;

Matches 581; Conservative 0; Mismatches 310; Indels 0; Gaps 0;

Db 122 TGATACCTCTTCAGAGTCATCGCTATGACCTAATCCACAGTCCAAAGT 181

627 tagtcatcttcgacacatgtgtgagatcaagaagcctcttctgtgtgccaacg 686

687 gtagcagcagccccctctatggaacagcaagaacagagctttgtggagatgctaacca 746

Db 242 GTGTTGCGGCCCCCTTTGTGGGACAGTAAGAAGCAGTGTPTTGTGGGATGCTGACCA 301

Db 302 TCACCGACTTCATCAGCATTTTGGACCGAATAAGTCAGCCCTGGTCAGATTACG 3611

807 agattgaacacataaagattgagacctggagagagatctaccctgcaggctgtcttcagc 866

DD 302 AACATCGAGGACGCACAAUAI AGAGACUGUGAAGAGAGUGIATCCUCGCGGAGCACCCTTTCGGC 769

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 11:32:25 ; Search time 5701.1 Seconds
(without alignments)
4076.714 Million cell updates/sec

Title: US-09-826-581-3
Perfect score: 1722
Sequence: 1 cctggcccccaagcaaga.....gatgagagcctcggcctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estc:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	148	8.6	572	10	B1344527 373008 MA
2	132.4	7.7	413	9	AA178898 zp38d10.r
3	116.8	6.8	1042	12	CNS03KJ0
4	82.6	4.8	933	10	BG919314 602817782
5	82.2	4.8	633	10	BC072114 BC072114
6	81.8	4.8	536	10	BM488662 pgm2n.pk0
7	80.2	4.7	595	10	BM487789 pgm2n.pk0
8	80.2	4.7	636	10	BC713637 pgl1n.pk0
9	80.2	4.7	647	10	BM440762 pgl1n.pk0
10	80.2	4.7	649	9	AJ395115 AJ395115
11	78.6	4.6	758	9	AJ396118 AJ396118
12	76.4	4.4	576	9	AV603335 AV603335
13	73	4.2	564	9	AV608257 AV608257
14	71	4.1	775	10	B1833269 603087149
15	70.4	4.1	448	9	AA558845 n169b09.s
16	70.4	4.1	450	10	BR351397 CM4-HR050
17	70.4	4.1	455	9	AA578219 n156h03.s

18	70.4	4.1	469	9	AL047390 DKFP586A
19	70.4	4.1	473	9	BE166881 CM4-HT050
20	70.4	4.1	508	9	BE166874 CM4-HT050
21	70.4	4.1	583	9	AW379936 RC4-HT025
22	70.4	4.1	586	10	BG609812 323791 MA
23	70.4	4.1	591	9	AA410926 fH09C01.x
24	70.4	4.1	598	9	AA418626 MR0-HT024
25	70.4	4.1	616	9	AW956906 EST368976
26	70.4	4.1	634	10	BC740148 602630747
27	70.4	4.1	668	10	BC705895 602669396
28	70.4	4.1	708	9	AW411228 fH11B02.x
29	70.4	4.1	710	10	BI223706 602943249
30	70.4	4.1	726	10	BI768590 603056916
31	70.4	4.1	742	10	BI914634 603179401
32	70.4	4.1	746	10	BF528081 602042828
33	70.4	4.1	756	10	BI859847 603386088
34	70.4	4.1	782	10	BI819312 603037761
35	70.4	4.1	782	10	BI821538 603038457
36	70.4	4.1	795	10	BE871189 601484767
37	70.4	4.1	808	10	BI771184 603054888
38	70.4	4.1	826	9	AL519198 AL519198
39	70.4	4.1	841	10	BI546514 603191524
40	70.4	4.1	864	10	BI858240 603384001
41	70.4	4.1	870	10	BI910928 603069326
42	70.4	4.1	903	10	BG177822 602314222
43	70.4	4.1	905	9	AL555228 AL555228
44	70.4	4.1	908	9	AL552459 AL552459
45	70.4	4.1	921	9	AL548987 AL548987

ALIGNMENTS

RESULT 1
LOCUS B1344527 572 bp mRNA linear EST 30-JUL-2001
DEFINITION B1344527 373008 MARC 2P1G Sus scrofa CDNA 5', mRNA sequence.
VERSION B1344527.1 GI:15037807
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 572)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized CDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross-match with the -minscore 18 and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 119 row: 1 column: 11
Seq primer: ATTTAGCTGACACTATG.
Location/Qualifiers
1..572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 2P1G"
/csize="type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

TITLE JOURNAL
COMMENT

FEATURES

source
1..572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 2P1G"
/csize="type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

OY 1298 ctggaccgggtgctcaggcaacgtactccacatccctcacaacacaacaacgcgtgctcaagttc 1357
 Db 256 attgatctccagctcacacagaaacccacttttaccattccacacacaaagagatctccaagttc 315
 OY 1358 ctggacatctcttgaag-----ccgggcccaggtggtggaggaaggggagactcg 1408
 Db 316 ctccacactgtttotgagctgcgacacatccccatccactgtgtcgaacaaagggccgttgaaat 375
 OY 1409 gcaagtgatcaagaggccctcagaagatctcagccctcagaagtcgtgggaagaagctggga 1468
 Db 376 ccacacctggtggggcgacagatcttgcgccatctgttccccatagtgcccaaaagccacaga 435
 OY 1469 gccctctgaagctctg-----gatccctgatctccacactggtcccaactcaacc 1520
 Db 436 cccctctgttttcggcccggaacacacccctgtcttacttttgccgctttgcctcgttttc 495
 OY 1521 agggctccctgctgcccgccctccctccctccacagcactatccaaatttgggacatcg 1580
 Db 496 acatctctgacgaacccaagcccccgttttatgacgacagaccctgacggagcttgggcaattgg 555
 OY 1581 gcaacatccagaaacttgacgtgctgctgctgaggaagacagacccaactcgaactgaactgaa 1640
 Db 556 gtacatctaccggacacattgctttcatctaccccgacacgcccattcatcaaaagcgcttaaca 615
 OY 1641 tcttttggagccgggtgctgctgacactgctgctgctgaagaatggtgtaccaccac 1698
 Db 616 tcttttggagaggggggggtgtgcgcttcctccgctggtgacacacatccgttaggaacccc 673

RESULT	4
LOCUS	BC919314
DEFINITION	BC919314 933 bp mRNA linear EST 05-JUN-2001 608017782P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:494661 5'
ACCESSION	BC919314
VERSION	BC919314
KEYWORDS	BC919314.1 GI:14299790 EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

lukavtsova; metazoa; chordata; craniata; vertebrata; euteleostomi;
mammalia; eutheria; rodentia; sciurognathi; muridae; murinae; mus
1 (bases 1 to 933)
nih-mgc <http://mgc.nhl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES
source
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10895 row: d column: 22
High quality sequence stop: 498.
Location/Qualifiers
1..933

BASE COUNT

Providing samples: Veriney Green, M.D., NIH

ORIGIN

Query Match	4.8%;	Score 82.6;	DB 10;	Length 933;
Best Local Similarity	52.1%;	Pred. No. 7.2e-08;		
Matches 232;	Conservative	0;	Mismatches 209;	Indels 4; Caps 2

OY	1242	gcgttttgaagcctgctctaaacctctaaagaagaccggtatccatgcctgcctgttcttg	1307
Db	7	gtccgggttgatgcgtctcttttcattatatttcgaaatattgaatccacagcgtctccagattatgc	66
OY	1302	accgcgtgtcagcgaacgctactccacatccctccacacaacgcgcctcccaagtlcttgc	1364
Db	67	accagagtcagcgaacaccttgtaactctcttactcacaagcggatcctcaagttcttca	126
OY	1362	acacattcttgaag---cctgtgcccaggtgtggaaggaaggggaagacctgtgscagttgatc	1418
Db	127	agttgttttgtaagtaaaccttcagccgcattatcccatattaccgtctgaaattccacacacacac	186
OY	1419	agaagagccttagagatctcttcagccctcagcagctgtcgtgggaagaagctcggagagccctctga	1478
Db	187	agacgcggttggggacagagagcggctgttggttgattatattgagcgttagacactcttcacatg	246
OY	1479	agctgtctgataccctcgtatctcaactcgtgtcccatccctaaaccagaggttccctcgtctgcc	1538
Db	247	ctctt-atcccgaaagacacctgcamaagccaattgcccccttctccagatnaccacggattccccca	305
OY	1539	ggccctctctctccctccagccagacatcccaagatttggatctgagcatalccgaagacttgg	1598
Db	306	agcccggaatttcattgtcttaactctctccaaagacctgcagatttgccacgtatggcaattatg	365
OY	1599	ctgtgtgtctgtagaagacagaccatccctgactgacatgacatctcttggacggcgctg	1658
Db	366	ccatagtcctctactaccacagccgctgctacgtgctgtggcgtactctttacagcacacggag	425
OY	1659	tgtctgacatgcctgtgtgtgtcaaga	1683
Db	426	tcttcgccttaccctgttagtgatga	450

RESULT	5
Bj072114	
LOCUS	633 bp mRNA linear EST J1-DEC-2001
DEFINITION	Bj072114 NIBB Mochii normalized Xenopus tailbud library xenopus
ACCESSION	laevis cDNA clone XL096j16 5', mRNA sequence.
VERSION	Bj072114
KEYWORDS	Bj072114.1 GI:17502303
SOURCE	EST.
ORGANISM	African clawed frog. Xenopus laevis

REFERENCE 1 (bases 1 to 633)

ALCAYALLA, A., TERASAKA, C., MOCHIL, M., UENO, N., SHIN-I, T. and KIHARA, Y.

TITLE	Expressed genes in <i>X. laevis</i> embryo
JOURNAL	Unpublished (2001)
COMMENT	Contact: Tadasu Shin-i

Center for Genetic Resource Information
National Institute of Genetics
1111 Yato, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
location/Qualifiers
1. 6333

```
/organism="Xenopus laevis"  
/db_xref="taxon:8355"  
/clone="XL096j16"  
/clone_lib="NIBB Mochi normalized Xenopus tailbud  
library"  
/tissue_type="whole embryo"
```

BASE COUNT 144 a 162 c 148 g 179 t
ORIGIN

Query Match 4.8%: Score 82.2; DB 10; Length 633;
Best Local Similarity 73.4%; Pred. No. 7.4e-08;
Matches 105; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Oy 1228 ctcccatctcgcagcctgttgaagctgtctacacccctcaagaacgcgataccatcg 1287
Db 10 CTCCTCCAGACGACAGCTTTTCCAGCTGTGATTCACATCAAGATAAGATCCACCG 69
Oy 1288 cctgcctgtcttgcacccggtgtcgaacgctacccatccatccacacaaacgcct 1347
Db 70 CCTACCAAGTATGATGATCCGATATCGCACACATCTTGATATCTCCATTAACGCCT 129
Oy 1348 gctcaagttcctgcacatcttgg 1370
Db 130 ACTCAAGTCTCTCACCTCTTTG 152

RESULT 6 536 bp mRNA linear EST 07-FEB-2002
LOCUS BM488662
DEFINITION pgm2n.pk008.g21 Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
clone pgm2n.pk008.g21 5' similar to gb|AAC52580.1 (U42413)
5'-AMP-activated protein kinase, gamma-1 subunit (Rattus norvegicus
1, mRNA sequence.
BM488662
VERSION BM488662.1 GI:18609593
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 536)
Cogburn, L.A. and Monsonego-Ornan, E.
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome
Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

JOURNAL COMMENT
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
source
1..536
Location/Qualifiers

/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk008.g21"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg: Embryo(d19): post-hatch(1d, 1, 3, 5, 7, 9
11 weeks); growth plate(1d, 7d, 14d post-hatch)"
/lab_host="E. coli EMDH108"
/note="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

BASE COUNT 117 a 171 c 132 g 116 t
ORIGIN

Query Match 4.8%: Score 81.8; DB 10; Length 536;
Best Local Similarity 70.1%; Pred. No. 8.4e-08;
Matches 110; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Oy 1228 ctcccatctcgcagcctgttgaagctgtctacacccctcaagaacgcgataccatcg 1287
Db 85 CTCCTCCAGACGACAGCTTTTCCAGCTGTGATTCACATCAAGATAAGATCCACCG 144
Oy 1288 cctgcctgtcttgcacccggtgtcgaacgctacccatccatccacacaaacgcct 1347
Db 145 CTCCTCCGTCATCGACCCCGACCTGGGCAACACTCTTATCATCTCCATCAACCAACGCAAT 204
Oy 1348 gctcaagttcctgcacatcttgg 1384
Db 205 CCTCAAGTCTCTCAAACTTTTATTCGACAGGTCCCA 241

RESULT 7 595 bp mRNA linear EST 07-FEB-2002
LOCUS BM487789
DEFINITION pgm2n.pk005.j24 Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
clone pgm2n.pk005.j24 5' similar to gb|AAC52580.1 (U42413)
5'-AMP-activated protein kinase, gamma-1 subunit (Rattus norvegicus
1, mRNA sequence.
BM487789
VERSION BM487789.1 GI:18608720
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 595)
Cogburn, L.A. and Monsonego-Ornan, E.
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome
Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

JOURNAL COMMENT
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
source
1..595
Location/Qualifiers

/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk005.j24"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg: Embryo(d19): post-hatch(1d, 1, 3, 5, 7, 9
11 weeks); growth plate(1d, 7d, 14d post-hatch)"
/lab_host="E. coli EMDH108"
/note="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

BASE COUNT 126 a 199 c 137 g 133 t
ORIGIN

Query Match 4.7%: Score 80.2; DB 10; Length 595;
Best Local Similarity 69.4%; Pred. No. 2e-07;
Matches 109; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1228 ctccacatctctgagcctgtgttgaagctgtctacaccctcatcaagaacggatccatcg 1287
LOCUS BM440762 647 bp mRNA linear EST 01-FEB-2002
DB 289 CTCGCCCAATGCGACGCTTTTGTATGCGCTCTCCGTGATCGGCAATGAATCCACCG 348
DEFINITION (pgrln.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
QY 1288 cctgcctgtctctgacccgggtgcagcaagcagctacccatccatccacacaaagcct 1347
pgrln.pk002.19 similar to g14506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1 non-catalytic subunit; AMP-activated, gamma 1 non-catalytic, gamma-1 (Homo sapiens)
DB 349 CTCGCCCGTATGACCGCCGACCTCGGCAACACTCTCTACATCTCTACCCCAACGCAAT 408
protein kinase, AMP-activated, gamma 1, mRNA sequence.
QY 1348 gctcaagttctcgcacatcttgtaagcctgggcca 1384
BM440762
DB 409 CCTCAAGTTCCTCAACTCTTTATGACAGGTCCCA 445
EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

ACCESSION BC713637 636 bp mRNA linear EST 08-MAY-2001
VERSION pgrln.pk008.c13 Normalized Liver Library Gallus gallus cDNA clone
KEYWORDS pgrln.pk008.c13 5' similar to g14506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1 non-catalytic subunit; AMP gamma 1; protein kinase, AMP-activated, noncatalytic, gamma-1 (Homo sapiens)
SOURCE 911273/489 ref|XP_006778.2| protein kinase, AMP-activated, gamma 1 non, mRNA sequence.
ACCESSION BC713637
VERSION HG713637.1 GI:14007587
KEYWORDS EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 636)
AUTHORS Burnside, J., Morgan, R.W. and Cogburn, L.A.
TITLE Chicken ESTs from a normalized liver library
JOURNAL Unpublished (2001)
COMMENT Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302 831-1345
Fax: 302-831-3411
Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES
source location/Qualifiers
1..636
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pgrln.pk008.c13"
/clone_lib="Normalized Liver Library"
/sex="Male and Female"
/tissue_type="liver"
/lab_host="E.coli EMDH10B"
/note="Vector: pCMVSPORT 6"

BASE COUNT 129 a 215 c 167 g 119 t 6 others
ORIGIN

Query Match 4.7%: Score 80.2; DB 10; Length 636;
Best Local Similarity 69.4%: Pred. No. 2.1e-07;
Matches 109; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1228 ctccacatctctgagcctgtgttgaagctgtctacaccctcatcaagaacggatccatcg 1287
LOCUS BM440762 647 bp mRNA linear EST 01-FEB-2002
DB 289 CTCGCCCAATGCGACGCTTTTGTATGCGCTCTCCGTGATCGGCAATGAATCCACCG 348
DEFINITION (pgrln.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
QY 1288 cctgcctgtctctgacccgggtgcagcaagcagctacccatccatccacacaaagcct 1347
pgrln.pk002.19 similar to g14506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1 non-catalytic subunit; AMP-activated, gamma 1 non-catalytic, gamma-1 (Homo sapiens)
DB 349 CTCGCCCGTATGACCGCCGACCTCGGCAACACTCTCTACATCTCTACCCCAACGCAAT 408
protein kinase, AMP-activated, gamma 1, mRNA sequence.
QY 1348 gctcaagttctcgcacatcttgtaagcctgggcca 1384
BM440762
DB 409 CCTCAAGTTCCTCAACTCTTTATGACAGGTCCCA 445
EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

RESULT 9
BM440762 647 bp mRNA linear EST 01-FEB-2002
LOCUS BM440762 647 bp mRNA linear EST 01-FEB-2002
DEFINITION (pgrln.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
QY 1288 cctgcctgtctctgacccgggtgcagcaagcagctacccatccatccacacaaagcct 1347
pgrln.pk002.19 similar to g14506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1 non-catalytic subunit; AMP-activated, gamma 1 non-catalytic, gamma-1 (Homo sapiens)
DB 349 CTCGCCCGTATGACCGCCGACCTCGGCAACACTCTCTACATCTCTACCCCAACGCAAT 408
protein kinase, AMP-activated, gamma 1, mRNA sequence.
QY 1348 gctcaagttctcgcacatcttgtaagcctgggcca 1384
BM440762
DB 409 CCTCAAGTTCCTCAACTCTTTATGACAGGTCCCA 445
EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 647)
AUTHORS Cogburn, L.A. and Nys, Y.
TITLE ESTs from Normalized Chicken Reproductive Tract cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Larry A. Cogburn
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES
source location/Qualifiers
1..647
/organism="Gallus gallus"
/strain="Commercial broiler and layer"
/db_xref="taxon:9031"
/clone="pgrln.pk002.19"
/clone_lib="Normalized Chicken Reproductive Tract cDNA library (pgrln)"
/sex="Male and Female"
/tissue_type="testis, ovary and oviduct"
/dev_stage="Various stages: embryonic, post-hatch, immature and sexually-mature"
/lab_host="E. coli EMDH10B"
/note="Vector: pCMVSPORT6; library made from three total RNA pools from each tissue (testis 25%, ovary 25%, and oviduct 50% of final RNA pool); Single pass sequencing from 5'-end"

BASE COUNT 137 a 222 c 160 g 128 t
ORIGIN

Query Match 4.7%: Score 80.2; DB 10; Length 647;
Best Local Similarity 69.4%: Pred. No. 2.1e-07;
Matches 109; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1228 ctccacatctctgagcctgtgttgaagctgtctacaccctcatcaagaacggatccatcg 1287
LOCUS BM440762 647 bp mRNA linear EST 01-FEB-2002
DB 289 CTCGCCCAATGCGACGCTTTTGTATGCGCTCTCCGTGATCGGCAATGAATCCACCG 348
DEFINITION (pgrln.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
QY 1288 cctgcctgtctctgacccgggtgcagcaagcagctacccatccatccacacaaagcct 1347
pgrln.pk002.19 similar to g14506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1 non-catalytic subunit; AMP-activated, gamma 1 non-catalytic, gamma-1 (Homo sapiens)
DB 349 CTCGCCCGTATGACCGCCGACCTCGGCAACACTCTCTACATCTCTACCCCAACGCAAT 408
protein kinase, AMP-activated, gamma 1, mRNA sequence.
QY 1348 gctcaagttctcgcacatcttgtaagcctgggcca 1384
BM440762
DB 409 CCTCAAGTTCCTCAACTCTTTATGACAGGTCCCA 445
EST.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

RESULT 10
AJ395115 649 bp mRNA linear EST 25-JAN-2001
LOCUS AJ395115 649 bp mRNA linear EST 25-JAN-2001
DEFINITION AJ395115 dkf426 Gallus gallus cDNA clone 21c2f1, mRNA sequence.
ACCESSION AJ395115
VERSION AJ395115.1 GI:7125706
KEYWORDS EST.


```

SOURCE      chicken.
ORGANISM    gallus gallus
REFERENCE   1 (bases 1 to 649)
AUTHORS     Abdrekhanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            J., Korn,B. and Buerstedde,J.M.
TITLE       A large database of chicken bursa ESTs as a resource for the
            analysis of vertebrate gene function
JOURNAL     Genome Res. 10 (12), 2062-2069 (2000)
MEDLINE     20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
            Location/Qualifiers
                source
                    1. 649
                        /organism="Gallus gallus"
                        /strain="CG"
                        /db_xref="taxon:9031"
                        /clone="21c271"
                        /clone_11b="dkfz426"
                        /tissue_type="Bursa of Fabricius"
BASE COUNT  130 a      227 c      155 g      137 t
ORIGIN
Query Match      4.7%; Score 80.2; DB 9; Length 649;
Best Local Similarity 69.4%; Pied. No. 2,1e-07;
Matches 109; Conservative 0; Mismatches 48; Indels 0; Gaps 0

Oy  1228  ctcccatcctcgagcctgttgaagcgtcttaccacccatcaagaacgcgatcatcg 1287
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db  46  ctcccccatgcatgcacgctttttgatccctctcctccgatccgcaatgaatccacg 105
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Oy  1288  cctgcctgtctcttgaccgggtgccaagcaagctactccacatcctcacacaacgcct 1347
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db  106  cctccccgcacatgcaccccgacactcgccgaacactctacatctccaccccaaacgcgat 165
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Oy  1348  gctcaagtcctgcacatcctltgtaagcctggccca 1384
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db  166  cctcaagttctccaacctctttattgacgaaggtccca 202
      |||||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

RESULT  11
AJ396118      758 bp      mRNA      linear      EST 25-JAN-2001
LOCUS       AJ396118 dkfz426 Gallus gallus cDNA clone 25f16r1, mRNA sequence.
DEFINITION AJ396118
ACCESSION  AJ396118
VERSION    AJ396118.1 GI:7127728
KEYWORDS   EST.
SOURCE     chicken.
ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 758)
            Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            J., Korn,B. and Buerstedde,J.M.
            A large database of chicken bursa ESTs as a resource for the
            analysis of vertebrate gene function
            Genome Res. 10 (12), 2062-2069 (2000)
            20568495
            Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
            Location/Qualifiers
                source
                    1. 758
                        /organism="Gallus gallus"

```

[illegible]

	Matches	101.	Conservative	0:	Mismatches	43:	Indels	0:	Gaps	0:
QY	1240	cagcctglltgaagctgtctacacctcaagaacggatccatcgctgcgtgtct								1299
Db	432	CAGCTTGTGTGATGCTGCTCTTCATTATATGGAACAAGATCCACAGCTCCAGTTAT								491
QY	1300	tgaaccggtgtgaagcaacgtactcacatctctacacacaaagcgtgctaaagtct								1359
Db	492	TCACCCGGAATCAGGCAACACCTGTATCATNCTCACCAAGGCCCATNCTCAAGTTCCT								551
QY	1360	gcacatctcttgaagcclggggccc								1383
Db	552	CAAGTTATTTTATCACCCGAGTTCCC								575

FEATURES	source
RESULT 13	
AV608257	
LOCUS	
DEFINITION	564 bp mRNA linear EST 28-NOV-2001
VERSION	AV608257 Bos taurus kidney fetus Bos taurus cDNA clone EIK1045H03
KEYWORDS	5', mRNA sequence.
SOURCE	AV608257.1 GI:9738630
ORGANISM	EST.
	cow.
	Bos taurus
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
	Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae;
	Bovidae: Bovinae: Bos.
	1 (bases 1 to 564)
	Takasuga,A., Hitotsune,S., Itoh,R., Jitchozono,A., Suzuki,H., Aso,H.,
	and Sugimoto,Y.
	Establishment of a high throughput EST sequencing system using
	poly(A) tail-removed cDNA libraries and determination of 36,000
	bovine ESTs
	Nucleic Acids Res. 29 (22), E108 (2001)
	21570554
JOURNAL	Contact: Yoshikazu Sugimoto
MEDLINE	Animal Genetics Division
COMMENT	Shirakawa, Institute of Animal Genetics
	Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
	Tel: 81-248-25-5641
	Fax: 81-248-25-5725
	Email: kazusugieccoc@ocn.ne.jp
	Single pass sequencing.
	This clone was obtained from a polyA-deleted cDNA library.
	Location/Qualifiers
	1..564

BASE COUNT	ORIGIN
142 a	145 c 121 g 155 t 1 others

Query Match	4.28;	Score 73;	DB 9;	Length 564;
Best Local Similarity	72.38;	Pred. NO. 7.7e-06;		
Matches	94;	Conservative	0;	Mismatches 36; Indels 0; Gaps 0;
QY 1240	cagcgcgtttgaagctgtctacacccctacgaagcagcatcgccgtcgtctct	1299		
DB 434	CAGCTGTTTGATGCTCTCTCTTCATTAAATTCGAAACAAAGATCCACAGGCTCCAGTTAT	493		
QY 1300	tgaaccggtgtcaggaacagtaactccacatctctacacacaaagcgtcgtcaagttcct	1359		
DB 494	TGACCGGATCAAGGCAACCGCTGTTCATCTCTACCCCAAGCGCATCTCTCAAGTTCT	553		
QY 1360	gcacacatcctt	1369		

```

      | | | |
Db 554 CAAGTTATT 563

```

RESULT	14
B1833269	
LOCUS	775 bp mRNA linear EST_04-OCT-2001
DEFINITION	603087149P1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5226132 5' ,
ACCESSION	mRNA sequence.
VERSION	B1833269
KEYWORDS	B1833269.1 GI:15944819
SOURCE	EST.
	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 775)	NIH-MGC	http://mgc.nci.nih.gov/	
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabbs@email.nih.gov			
	Tissue Procurement: Life Technologies, Inc.			
	CDNA Library Preparation: Life Technologies, Inc.			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLM1569	row: a	column: 13	
	High quality sequence stop: 773.			

```

FEATURES
  source
    Location/Qualifiers
      1..775
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5226132"
        /clone_11b="NIH_MGC_120"
        /lab_host="DH10B"
        /note="Organ: pooled pancreas and spleen; Vector:
        pCW-SpOrf6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
        source anonymous pool of spleen and pancreas from 28 yo
        male. Library is oligo-dT primed and directionally cloned
        (EcoRV) site is destroyed upon cloning). Average insert
        size 1.5 kb, insert size range 1-2.5 kb. Library is
        normalized and enriched for full-length clones and was
        constructed by C. Gruber (Invitrogen). Research Genetics
        tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT
  200 a 194 c 171 g 209 t 1 others

```

Query Match	Similarity	68.1%	Score 71	DB 10	Length 775
Best Local	Similarity	68.1%	Pred. No. 2.5e-05		
Matches	99	Conservative	0	Mismatches 46	Indels 0
				Gaps 0	
QY	1240	cagccgtgttgaagctgtctacacccctacaaagacggaatccatgcgcctgtct	1299		
Db	455	CAGCTGTTTGATGCTCTCTCTTCATTATTCGGAACAAGATTCACAGCTGCAGTTAT	514		
QY	1300	tgaacgcggtgtcaagcaagtaactccaatcttcacacaaagacgctgtcattcct	1359		
Db	515	TGACCCGAAATCAGGCAATCTTTGTACATCTCCACCAAGGCAATTCGAAAGTTCT	574		
QY	1360	gcacatcttgttaagcctggccc	1383		
Db	575	CAAAATGTTTATCACTGATGTTCC	598		

RESULT	15				
AA558845/c					
LOCUS					
DEFINITION	AA558845	448 bp	mRNA	linear	EST 09-SEP-1997
	n165b09.s1				
	similar to SW:AAKC_RAT P80385 5'-AMP-ACTIVATED PROTEIN KINASE,				
	IMAGE:1045913				

CAMMA CHAIN ; mRNA sequence.
AA558845

ACCESSION	AA558845	
VERSION	AA558845.1	GI:2329612
KEYWORDS		

SOURCE	human.
human.	human.

ORGANISM

REFERENCE

AUTHORS
TITLE

-
-
-
-
-

JOURNAL
COMMENT

SOURCE

BASE COUNT

ORIGIN

Query Mat

Best Local Matchbox

References

1240

Db 238 C

QY 1300 F

Db 178 T

QY 1360 g

Db 11.8 C

Job time: 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 11:28:00 ; Search time 7316.32 Seconds

(without alignments)
4925.358 Million cell updates/sec

Title: US-09-826-581-3
Sequence: 1 ccctggcccccagatcaaga.....gagagagagctcgagctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pal: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_pl: *
25: em_pi: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	1722	100.0	1722	6	AX281580	AX281580 Sequence
C	2	1719.4	99.8	152129	2	AC027416 Homo sapi
C	3	1697	98.5	196554	2	AC073128 Homo sapi
C	4	1697	98.5	206854	4	AC009974 Homo sapi
C	5	612.6	35.6	5888	4	AF214521 Sus scrofa
C	6	209.4	12.2	227224	2	AF336381 Mus muscu
C	7	168.2	9.8	1647	6	AX281582 Sequence
C	8	168.2	9.8	2109	6	AX099776 Sequence
C	9	168.2	9.8	2115	6	AX099802 Sequence
C	10	168.2	9.8	2115	6	AF214519 Sequence
C	11	168.2	9.8	2290	6	HS249977 Sequence
C	12	148	8.6	1867	6	AX099774 Sequence
C	13	148	8.6	1873	4	AX214520 Sus scrofa
C	14	148	8.6	1873	6	AX099800 Sequence
C	15	148	8.6	2022	6	AX099804 Sequence
C	16	91.4	5.3	106	11	G67375 Human
C	17	82.2	4.8	142903	2	AL627254 Drosophila
C	18	79.4	4.6	14411	4	AF329081 Bos taurus
C	19	77	4.5	73638	2	AC015613 Homo sapi
C	20	74.2	4.3	81210	2	AC019242 Homo sapi
C	21	74.2	4.3	171822	2	AC011603 Homo sapi
C	22	74.2	4.3	180749	2	AC025256 Homo sapi
C	23	70.4	4.1	1578	9	HS042412 Homo sapi
C	24	70.4	4.1	1677	9	BC000358 Homo sapi
C	25	69.4	4.0	101215	9	AC006966 Homo sapi
C	26	68	3.9	3497	3	AF094763 Drosophila
C	27	68	3.9	80069	2	AC019671 Drosophila
C	28	68	3.9	173634	3	AC009344 Drosophila
C	29	68	3.9	195868	3	AC008308 Drosophila
C	30	68	3.9	230266	3	AE003733 Drosophila
C	31	67.6	3.9	7218	6	166494 Sequence
C	32	65.8	3.8	2082	3	AF094764 Drosophila
C	33	65.8	3.8	4666	3	AY070541 Drosophila
C	34	65.6	3.8	104	10	AF266982 Rattus norvegicus
C	35	65.6	3.8	1328	10	RNMPKGM
C	36	65.6	3.8	1550	10	RNMPKGM
C	37	65.6	3.8	1623	10	AF036535 Mus muscu
C	38	63.8	3.7	1167	9	AB025580 Homo sapi
C	39	63.8	3.7	1435	6	AR139104 Sequence
C	40	63.8	3.7	2062	9	AF249976 Homo sapi
C	41	63.8	3.7	2194	9	AF087875 Homo sapi
C	42	63.8	3.7	2203	9	BC020540 Homo sapi
C	43	63.8	3.7	2223	9	AK001887 Homo sapi
C	44	63.8	3.7	3132	10	BC015283 Mus muscu
C	45	62.4	3.6	125020	9	AF429315 Homo sapi

ALIGNMENTS

RESULT	1	AX281580	1722 bp	DNA	linear	PAT 03-NOV-2001
LOCUS	AX281580	Sequence 3 from Patent WO0177305.				
DEFINITION	AX281580					
ACCESSION	AX281580					
VERSION	AX281580.1	GI:16608831				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS		Andersson, L., Luthman, H. and Marklund, S.				
TITLE		Variants of the human amp-activated protein kinase gamma 3 subunit				
JOURNAL		Patent: WO 0177305-A 3 18-Oct-2001;				
AREXIS AB (SE)		Location/Qualifiers				
FEATURES		1..1722				
SOURCE		/organism="Homo sapiens"				
BASE COUNT		321 a 504 c 534 g 363 t				
ORIGIN						

misc_feature	24420. .27059
misc_feature	/note="assembly-fragment"
misc_feature	27160. .30170
misc_feature	/note="assembly-fragment"
misc_feature	30271. .33968
misc_feature	/note="assembly-fragment"
misc_feature	34069. .38179
misc_feature	/note="assembly-fragment"
misc_feature	38280. .42366
misc_feature	/note="assembly-fragment"
misc_feature	42467. .46365
misc_feature	/note="assembly-fragment"
misc_feature	46466. .51285
misc_feature	/note="assembly-fragment"
misc_feature	51386. .55871
misc_feature	/note="assembly-fragment"
misc_feature	55972. .60595
misc_feature	/note="assembly-fragment"
misc_feature	60696. .66595
misc_feature	/note="assembly-fragment"
misc_feature	66596. .73218
misc_feature	/note="assembly-fragment"
misc_feature	73319. .77115
misc_feature	/note="assembly-fragment"
misc_feature	77216. .85022
misc_feature	/note="assembly-fragment"
misc_feature	85123. .93314
misc_feature	/note="assembly-fragment"
misc_feature	93415. .101193
misc_feature	/note="assembly-fragment"
misc_feature	101294. .113090
misc_feature	/note="assembly-fragment"

Query Match	99.8%	Score 1719.4;	DB 2;	Length 152129;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1720:	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	1	ccctgagccctcagatcagaagaagcctctctctgctcgtgtgccaagctgtgcggagcagc	60
Db	36618	ccctgagccctcagatcagaagaagcctctctctgctcgtgtgccaagcgtgtcgggacac	36559
OY	61	ccctcatalgagcagcaagaagcagacgtcttgaggtgagagagctgaggaagctgaag	120
Db	36558	ccctctatgagcagcagcaagaagcagacgtcttgaggtgagagagctgaggaagctgaag	36499
OY	121	ggagatggaaggaaggttgaaggaggagatctgtacggctgtctcctggagctgaccttgatat	180
Db	36498	ggagatggaaggaaggttgaaggaggagatctgtgtacggctgtctcctggagctgaccttgatat	36439
OY	181	accacaagcttctgctctcagagcccaagcccaagcagagggccagagggtgagagaagctcatcc	240
Db	36438	accacaagcttctgctctcagagcccaagcccaagcagagggccagagggctgagagaagctcatcc	36379
OY	241	ggagctctgcaatgagccagcttgagagaccctggagctcaalttccccaatctgtagagccgct	300
Db	36378	ggagctctcagatgagccagcttgagagaccctggagctcaalttccccaatctgtagagccgct	36319
OY	301	atgacacagctgacacacttccacctcgcgtacatcgatctgagccctgtgacatagtgtaagg	360
Db	36318	atgacacagctgacacacttccacctcgcgtacatcgatctgagccctgtgacatagtgtaagg	36259
OY	361	agcaaaatgggggggagagcagagagaagaagaccaccaactctcaagccttgagggtctgcccc	420
Db	36258	agcaaaatgggggggagagcagagagaagaagaccaccaactctcaagccttgagggtctgcccc	36199
OY	421	actgtccgttcccaacagctccccaactgtgtctctcagacacaaaggaagaacttgcaaggttgagga	480
Db	36198	actgtccgttcccaacagctccccaactgtgtctctcagacacaaaggaagaacttgcaaggttgagga	36139
OY	481	ggggaatctgaccctcaaccctgctcttccaccaagaagcccgaggcttgacactctccctccgac	540
Db	36138	ggggaatctgaccctcaaccctgctcttccaccaagaagcccgaggcttgacactctccctccgac	36079

QY	541	ccccccctgaggaatgctcgacacatactgacttaactctgtgtgctgcatgctactaca	600
Db	36078	CCCTCCCTCGAGGAGTGCAGACATACGACTTCATCTGTGGTGCATCTCGTACTTAACA	36015
QY	601	gtccccccctgaggaagctgagcttggaattctta-cgggacaccgaagggcggggcgg	660
Db	36018	GGTCCCCCGGTGGAGAGTGGGCTTGGAATCTTATGGCAGCCAAAGGGCCGGGGCGG	35955
QY	661	aagggaagctctctcgagcctgagctggtgccttagaagcccaagctctctcgattcgatgc	720
Db	35958	AGGGGAGTCTCTCTGTGAGGCTGTGTCCTCCATGAAGCCCACTTTCTGTGACTTCGAGATC	35895
QY	721	ctgtgattctcttaagttcacatactatagattggaacaacataaagattggaacctggag	780
Db	35898	CTGTCCATGTCTCTAGGTCACAGTCTATAGATTGAACACATTAATTGAGACTGGAG	35833
QY	781	gggtgagtgagggaaggaaccgggaagggcgctgtgtgtatgtgtggtgcgaaggtctaa	840
Db	35838	GGGTGAGTGGGGAGAGAAACCCGGAAAGGGGCTGTGTGATGTGGCCAGCGCTTAA	35775
QY	841	gttgagaagtatggagcagtgaggaatgtctcttgagt-ggaacaaggggaggaataagaagctcg	900
Db	35778	GTGAGAGATTGGGACGTGGGGATGTCTTGAGTGAACAGGGGAAGCAATAGGAGCTTG	35715
QY	901	ggtgcctgacggaaagggaagctgtgcttggagcttgacaagtgaagcgaaigtgaccgctcccc	960
Db	35718	GGTGCCTGACGGAAAGGAAGCTGCCGGGACATGCAAGGTGAGGCAAGTACCGGCTCCCC	35658
QY	961	tggcgtgactcttgagcctctcttcgcaagaattcaacgcgaagctgtgttaagccttggt	1020
Db	35658	TGGCTGACTCTGGCTCTTTTGTGCAGAGATCTACCTGCACAAAGCTGTTAAGCTCTGGT	35598
QY	1021	ctccatctctctctaatagtatagtggtgtgtctgtcattcaatcaccttgagcctctctccc	1080
Db	35598	CTCCATCTCTCTCTTAATGATAGGGGGTGTCTGTGCTCATTTCACTGAGCTCTCTCCC	35538
QY	1081	aaagtcctccctcccaagctcccaactaaagctctgaaactcaacctcttaacttaagcggacaa	1140
Db	35538	AAAGTCTCCCTCCCAAGTCCCACTGACTCTTAACCTCAACCTTTATCTTACGGCGACA	35479
QY	1141	cagacaagaagacgttgtgctgcctcgccctctttaaaggccttggaatgaggtgtgtct	1200
Db	35478	CAGACAAGGAGACTTGGTGGCTTGCCTGCCCTCTTTTAAAGGGCTTGGAATGAGATTGTCT	35419
QY	1201	ctcccttaggcgtgcccgaggtcgaactgtccacatctctgcaagcgtgttgaagctgtcta	1260
Db	35418	CTCCCTAGGCTCTCCCGAGGCTACAGTGCCTCATCTCTCAACCTGTTTGAAGCTGTCTA	35359
QY	1261	caaccctcaacaagaaccggaatccatgcctctgctgttcttgaccgggtgtcaaggcaagt	1320
Db	35358	CACCTCTATCAAGAACCGGATTCATGTGCTGTCTGTTTGAACCCGGTCTCAGCAACT	35299
QY	1321	actccaatctctcacacacaaaagcgtctgcaagttctctgcacatcttltglaagcctgg	1380
Db	35298	ACTCCACATCTCTACACACAAACGGCTGCTCAAGTTTCGTCGACATCTTTGTAAGCTGG	35239
QY	1381	ccccagtgagggaagggggggaaccttggtgcaggtgatacaaggagcctgaaggcttcaag	1440
Db	35238	CCCCAGTGGGAGGAAGGGGAAACCTGTGGCAGGTATAGAGAGGCTCTGAGGAGCTTTCAG	35179
QY	1441	cccttagcaatctgttgaggaaagacttggaagccctcttgaaagctgagctggaatccctgaatccc	1500
Db	35178	CCCTAGCAATCTGTGGGGAAAGACTTGGAGCCCTTGTGAAGCTGCTGGATCCCTGATCTCC	35119
QY	1501	accgtgtgcccaatcctaaaccaagtgctcctgtgcctgcggccctctctctctaaacgac	1560
Db	35118	ACCTGTGTCCTCTCTTAACCAAGGTTTCCCTGTGCCCGGCGCTCTCTCTCTACGCAC	35059
QY	1561	tatccaagaatttgggcatctggacaactcttcagagacttggtctgtgtgcctgggaagaacc	1620
Db	35058	TATCCAAGATTTGGGCAATCGGCACATTCGGAAGACTTGGCTGTGGTGTGGAAGACAGCAC	34999
QY	1621	catctctgactgacatcgagacattttgtgaccgggtgtgtctctgcaactgctgtgtctaa	1680

|||||
Db 34998 CATCTACTGACATGACATCTTTTCGACCCGCGTGTCTGCACTGCTGTGCA 34939
Qy 1681 cgaatgtgtaaccacccacgaatgagagctcgagctg 1721
Db 34938 CGAATGTGTAACCACCCACGATGACAGCTCGGCTCG 34898
RESULT 3
AC073128/c 196554 bp DNA linear HTG 21-FEB-2001
DEFINITION Homo sapiens chromosome 2 clone RP11-64705, WORKING DRAFT SEQUENCE,
17 unordered pieces.
AC073128
AC073128.3 GI:13027579
HTG: HTGS_PHASE1: HTGS_DRAFT: HTGS_FULLTOP.
human.
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 196554)
JOURNAL Waterston,R.H.
AUTHORS The sequence of Homo sapiens clone
TITL 2 (bases 1 to 196554)
JOURNAL Waterston,R.H.
REFERENCE Direct Submission
TITL Submitted (08-JUN-2000) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Feb 21, 2001 this sequence version replaced gi:8469048.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0647005
----- Summary Statistics -----
Sequencing vector: MJ3; 988
Sequencing vector: plasmid; 08
Chemistry: Dye-terminator Big Dye; 08 of reads
Chemistry: Dye-terminator Big Dye; 08 of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 18705 bases at least Q40
Consensus quality: 190513 bases at least Q30
Consensus quality: 192099 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 194954; sum-of-ctigs
Quality coverage: 5.58 in Q20 bases; agarose-fp
Quality coverage: 5.67 in Q20 bases; sum-of-ctigs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1157: contig of 1157 bp in length
* 1158 1257: gap of unknown length
* 1258 3600: contig of 2343 bp in length
* 3601 3700: gap of unknown length
* 3701 5103: contig of 1403 bp in length
* 5104 5203: gap of unknown length
* 5204 8524: contig of 3321 bp in length
* 8525 8624: gap of unknown length
* 8625 11856: contig of 3232 bp in length
* 11857 11956: gap of unknown length
* 11957 15783: contig of 3827 bp in length
* 15784 15883: gap of unknown length
* 15884 21906: contig of 6023 bp in length

* 21907 22006: gap of unknown length
* 22007 28887: contig of 6881 bp in length
* 28888 28987: gap of unknown length
* 28988 35255: contig of 6268 bp in length
* 35256 35356: gap of unknown length
* 35356 44642: contig of 9287 bp in length
* 44643 44742: gap of unknown length
* 44743 58275: contig of 13533 bp in length
* 58276 58375: gap of unknown length
* 58376 73817: contig of 15441 bp in length
* 73817 73916: gap of unknown length
* 73917 92140: contig of 18224 bp in length
* 92141 92240: gap of unknown length
* 92241 113337: contig of 21097 bp in length
* 113338 113437: gap of unknown length
* 113438 130325: contig of 16888 bp in length
* 130326 130425: gap of unknown length
* 130426 149287: contig of 18662 bp in length
* 149288 149387: gap of unknown length
* 149388 149388: contig of 47167 bp in length.
* Location/Qualifiers
1. 196554
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-64705"
1. 1157
misc_feature
/note="assembly_name:Contig17"
1258 3600
/note="assembly_name:Contig18"
3701 .5103
/note="assembly_name:Contig19
clone_end:77
vector_side:right"
5204 .8524
/note="assembly_name:Contig20"
8625 .11856
/note="assembly_name:Contig21"
11957 .15783
/note="assembly_name:Contig22"
15884 .21906
/note="assembly_name:Contig23"
22007 .28887
/note="assembly_name:Contig24"
28988 .35255
/note="assembly_name:Contig25"
35356 .44642
/note="assembly_name:Contig26"
44743 .58275
/note="assembly_name:Contig27"
58376 .73816
/note="assembly_name:Contig28"
73917 .92140
/note="assembly_name:Contig29"
92241 .11337
/note="assembly_name:Contig30"
113438 .130325
/note="assembly_name:Contig31"
130426 .149287
/note="assembly_name:Contig32"
149388 .196554
/note="assembly_name:Contig33
clone_end:56
vector_side:right"

BASE COUNT 52296 a 46993 c 45889 g 49770 t 1606 others
ORIGIN

Query Match 98.5% Score 1697; DB 2: Length 196554;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1719; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 1 cctggccctcagatcaagaagccttctgtgctgtgccaacggtgtcgcgagc 60

|||||
Db 61117 CCTGGCCCTCAGATCAAGAGCCCTTCTTCTGCTGCGCAACGGGTGGGACAC 61058
Oy 61 cccctataaggacagcaagaagcagagctctg tgggtgagagagcctgaggagtgtaag 120
Db 61057 CCTCTATGGAGACAGCAAGAACAGAGCTTTGTGGGTAGAGAGAGCTGGAGAGGTGAAG 60998
Oy 121 ggaagatgagagagatgagagggagatctgtacaggtgtgtctgggctgagctctgatat 180
Db 60997 GCAGATGGAGAGCTGACGGCGAGATCTTGTACGCTTGTCTGGGGCTGATCTGATAT 60938
Oy 181 accacaagcttggtctcagaagcccaagccaagggccaaggttgaggaagatccatcc 240
Db 60937 ACCACAAGCTTGGCTTCAAGGCCAGCCAGGGGCGAGGGTGGAGAAAGTCATCC 60878
Oy 241 ggaagctgagcagcagctgagagacccttgaggctcaatctccacatctgtgagagcgt 300
Db 60877 GCAGTCTGATGGCAGCTGGGAGACCCCTGGGGCTCAATTTTCCCATCTGTGAGACCGCT 60818
Oy 301 atgacacagctgagacaccttcaacctcagctactgcaatgagcctgtgcataggtgtaagg 360
Db 60817 ATGACACAGCTGACACCTTTCACCTCCCTACTGATGGCCCTGTCCATAGCTCTAGCG 60758
Oy 361 agcaaaatgaggagagcagagagaagaagagcccaactctcaaggtctgagggtgctgccc 420
Db 60757 AGCAAAATGGGGGAGGAGAGAGAAAGAGCCCACTTCTAGGCTGGGGGCTGCCCC 60698
Oy 421 acgtgccgtgtcccaagctcccaactgtgtctcagcaagaagacatgagcaggttgaggga 480
Db 60697 ACTGTCTCTTCCACACTCCCACTGTGTCTCAGCACAGAGACTGGAGGGTGGGGA 60638
Oy 481 gggagatctgaaacctcaacctgtctccaaagagcccgaggcagactctctcccgcc 540
Db 60637 GGGATCTGACCTCAACTGCTTCCACCAAGAGCCCGGGGCTGACCTCTCCCGCC 60578
Oy 541 cctccctcagcagagatgctgacacatcaactgactatccctgtgtctgacgtactaca 600
Db 60577 CCT-CCCTGCAAGGAGTGTGACCATCACTGACTTCTGCTGCTGCTGCTGCTACTACA 60519
Oy 601 ggtcccccctgtgagaggtgagcgtgggaatcttaaggagagcccaagagggcgaggcg 660
Db 60518 GGTCCCCCTGTGTGAGAGTGGGAGTGGGGAATCTTATGGGCAAGAGGGGCGGGGGCG 60459
Oy 661 agggagagctcctcagagcctgagcctgagcccaagcccaagctctctcagactctggaatc 720
Db 60458 AUGGAGTCTCTCTGACCTGAGCTGTGCTTACAGCCCAAGCTTCTTCTGAGATC 60399
Oy 721 ctgtcgaatgtctcaggtcagacataatgaaatgaaacataagaltgagacctgag 780
Db 60398 CTGTGATGTCTCTAGTTCAGATCTATGAGATTGACACATTAAGATTGAGACCTGGAG 60339
Oy 781 gggatgagtgaggagaggaacccggaagggctgtgtgtatgtgtggcgagggcttaag 840
Db 60338 GGGTATGAGGGGAGAGAGAACCCGGGAA-GGGCTGTGTGATGGTGGGCGAGGGCTTAAG 60280
Oy 841 gtgagagatgagcagctgaggatgtctcagagtgaaacagagggagaaacataagagcctcg 900
Db 60279 GTGGAGGATGGGCAAGTGGGATGTCTGAGTGAACAGGGAGGAGCAATAGAGCGCTCG 60220
Oy 901 ggtgctcagcagaaaggaagctgctgaggaactgaaaggtgagagcaggtgagcgcctccc 960
Db 60219 GGTGCTTACGGAAGAGGAGCTGCTGGGACTGCAAGGTGAGGAGTGAACCGGCTCCCC 60160
Oy 961 tggcctgagcttgctctctctcagagatctacatgcaagcgtcgtctcaagcctctggt 1020
Db 60159 TGGCCTGACTCTGGCTCTTCTGAGAGATCTACTGCAAGGCTGCTTCAAGCCTCTGCT 60100
Oy 1021 ctccatctctcctaataagatgaggtgtctctgtcattcaacctgagcctcctcctccc 1080
Db 60099 CTCCATCTCTCTAATAGTGGGTCTCTCTCTCATTCACCTGAGCCCTCCCTCC 60040
Oy 1081 aaagctccctcccaagctcccaactcagactcgaactcaacctctcatcctaggaggaca 1140
|||||

Db 60039 ACAGTCCCTTCCAGTCCCACTGACTGTGAACACTCTTCAATCTAGCGCGACA 59980
Oy 1141 cagaaagagagccttggtgcccgtccctctctttttagggccctggagatgaggttct 1200
Db 59979 CAGACAAGGAGCCTTGGTGGCTGCTCCCTCTTTTAAAGGCGCTGGGATGGAGGCTGTCT 59920
Oy 1201 ctccctaggctgcccagagctacatgctcctccatctcagcagcctgltttagagctgcta 1260
Db 59919 CTCCCTAGCTGCCCCGAGGCTCACTGCTCCCATCTGTGAGAGCCTGTGAACTGTCTA 59860
Oy 1261 caccctataaagaacggatccatgcctgctgtctctgtatgagccggtgtcaggaacgt 1320
Db 59859 CACCTCTATCAAGAACCGATCCATCGCTTCTGTTTCAACCGGTGTCAAGCAAGT 59800
Oy 1321 actccacatctctcaacaacaagcctgctcagatctcagatctcctgtttagagcctgag 1380
Db 59799 ACTCCACATCTCTCAACACAAGCCCTGCTCAAGTCTCTGACACTTTTGTAACTGCG 59740
Oy 1381 cccaggtgagagagagggagaccttgaggcaggtgatalcaagaaggcctgagaggtctcag 1440
Db 59739 CCCAGGTGGAGAGGAGGAGAGACTGGGAGCTGGGAGCTGATCAGAGGCTGAGAGAGTCTCAG 59680
Oy 1441 cccatgaggtgtgaggaagagctgagagccctcttgaagctgtgtatccctgatatcc 1500
Db 59679 CCTTAGCAGATGTGGGGAAGAGCTGGGAGCCCTTGTGAAGCTGTGATCCCTGATCTCC 59620
Oy 1501 accgtgtcccatcctcaacagaggttccctgtcgtcccgccctctcctcaacagac 1560
Db 59619 ACCTGTGCTCCATCTTAACAGAGGTTCCCTGCTGCTCCCGGCCCTCTTCTTACCCGAC 59560
Oy 1561 tatccaagatttgagcatcgagacatcccgagacttgagctgtgtctgagagagacac 1620
Db 59559 TATCCAAGATTGGGCAATCGGCACATTCGAGACTTGTGCTGTGCTGAGAGACGACC 59500
Oy 1621 catcctgacgtgacatgagacatcttctgagacggcggtgtctcagactgctgtgtcaa 1680
Db 59499 CATCTGACTGACTGGAACATCTTTGTGAGACGGCGTGTGTGCACTGCTGTGTCAA 59440
Oy 1681 cgaatgtgtaecccaaccccaagatgagaggtcctgagctg 1721
Db 59439 CGAATGTGTACCCACCCAGATGAGAGGCTGGGCTGG 59399

RESULT 4
AC009974/c 206854 bp DNA linear PRI 09-JAN-2002
LOCUS AC009974
DEFINITION Homo sapiens BAC clone RP11-459119 from 2, complete sequence.
ACCESSION AC009974
VERSION AC009974.9 GI:16799058
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 206854)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 206854)
AUTHORS Harris, A. and Colton, M.
TITLE The sequence of Homo sapiens BAC clone RP11-459119
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 206854)
AUTHORS Waterston, R.H.
TITLE Direct Submission
REFERENCE 4 (bases 1 to 206854)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1999) Genome Sequencing Center, Washington
MO 63108, USA
4 (bases 1 to 206854)
Submitted (08-NOV-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
5 (bases 1 to 206854)
AUTHORS
Waterston,R.H.
TITLE
Direct Submission
JOURNAL
Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
6 (bases 1 to 206854)
AUTHORS
Waterston,R.
TITLE
Direct Submission
JOURNAL
Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 8, 2001 this sequence version replaced gi:13431203.

COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0459119

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPI1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Calanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-1077K22: the clone sequenced to the right is RP11-64705. Actual start of this clone is at base position 1 of RP11-459119; actual end is at base position 206854 of RP11-459119.

Data from AC079810 and AC073128 was used to finish this clone, AC009974. Polymorphisms have been identified between AC073128 and AC009974. A single plasmid region exists between 38812-38903. An unresolved tandem in the HERV SVA exists between 184390-185163. PCR suggests that approximately 1700 bps are missing.

FEATURES
source
1..206854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-459119"
/clone_11b="RPI1-11"
1..37
/note="match to EST BE314060 (NID:g9134719)"

misc_feature
1..37
/note="match to EST BE314060 (NID:g9134719)"

misc_feature
1..37
/note="match to EST B1059713 (NID:g14467240)"

misc_feature
1..37
/note="match to EST BF183086 (NID:g11061273)"

misc_feature
1..37
/note="match to EST AL567345 (NID:g12920610)"

misc_feature
1..37
/note="match to EST AW880850 (NID:g8042860)"

misc_feature
1..37
/note="match to EST BF304755 (NID:g11251653)"

misc_feature
1..37
/note="similar to Homo sapiens EST B111348 (NID:g14565249)"

misc_feature
1..37
/note="match to EST BG477625 (NID:g13409904)"

misc_feature
1..37
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature
3..37
/note="match to EST BE908408 (NID:g10402954)"

misc_feature
4..37
/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

misc_feature
164..662
/note="match to EST C05773 (NID:g1502549)"

misc_feature
220..221
/note="match to EST BG470047 (NID:g13402322)"

misc_feature
281..344
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature
281..344
/note="match to EST BE908408 (NID:g10402954)"

misc_feature
281..344
/note="match to EST BF304755 (NID:g11251653)"

misc_feature
281..344
/note="match to EST BF183086 (NID:g11061273)"

misc_feature
281..344
/note="similar to Homo sapiens EST B111348 (NID:g14565249)"

misc_feature
281..344
/note="match to EST BG477625 (NID:g13409904)"

misc_feature
281..344
/note="match to EST B1059713 (NID:g14467240)"

misc_feature
281..344
/note="match to EST AL567345 (NID:g12920610)"

misc_feature
281..344
/note="match to EST BE314060 (NID:g9134719)"

misc_feature
281..344
/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

misc_feature
281..344
/note="match to EST AW880850 (NID:g8042860)"

misc_feature
294..344
/note="match to EST BG470047 (NID:g13402322)"

misc_feature
complement(357..358)
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"

misc_feature
588..929
/note="match to EST T64073 (NID:g667938) yc05d12.r1"

misc_feature
594..763
/note="match to EST BF304755 (NID:g11251653)"

misc_feature
594..763
/note="match to EST B1059713 (NID:g14467240)"

misc_feature
594..763
/note="match to EST BG477625 (NID:g13409904)"

misc_feature
594..742
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature
594..763
/note="match to EST BE908408 (NID:g10402954)"

misc_feature
594..763
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"

misc_feature /note="match to EST AI670836 (NID:g4850567)" wa04g10.x1"
594. .763
misc_feature /note="match to EST BF183086 (NID:g11061273)"
594. .763
misc_feature /note="similar to Homo sapiens EST B1114348
(NID:g14565249)"
599. .763
misc_feature /note="match to EST AL567345 (NID:g12920610)"
617. .1084
misc_feature /note="match to EST AA481361 (NID:g2210913)" zv44e01.r1"
622. .763
misc_feature /note="match to EST AI860958 (NID:g5514574)" w156f05.x1"
622. .763
misc_feature /note="similar to Mus musculus EST AI196847 (NID:g3749453)
w167e01.x1"
684. .763
misc_feature /note="match to EST BC992568 (NID:g14396638)"
962. .1084
misc_feature /note="match to EST AI656812 (NID:g4740791)" tt54b06.x1"
967. .1084
misc_feature /note="match to EST BE908408 (NID:g10402954)"
967. .1085
misc_feature /note="match to EST BF304755 (NID:g11251653)"
967. .1091
misc_feature /note="match to EST AA043371 (NID:g1521226)" zk53e10.r1"
967. .1071

Query Match	98.5%;	Score 1697;	DB 9;	Length 206854;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1719;	Conservative	0;	Mismatches	0;
			Indels	2;
			Gaps	2;

QY	1	ccctgagccctcagatcacaagaagccctctctgtctcgtgtgccaagcgtgtgcggcagc	60
Db	166697	CTGTGCCCCCTCGAGATCAAGAGGCCCTTCTTTTGCTCTGTGTGCCCAACGCTGTGCGGCAGC	16638
QY	61	ccctctatggagacagcaagaagcagaagctctgtgtgtgtgaagagagccttgggggagtgtaag	120
Db	166637	CCCTCTATGTGGGACAGCAAGAAGCAGACTTGTGTGGTGAGAGAGCGCTGGGAGGTGAAC	166578
QY	121	ggaga tggagagagagtggaagggggagatccttgtaagcgttctgtctgggcctgaacctctgatat	180
Db	166577	GCAGATGGAGGGAGGGAGGGGAGATCTTTGACGGTTGTTCTGGGGGTGATCTCTGTATA	166518
QY	181	accacaagcctctggtcttcagacaaagcccaagcagaaggccaaagctgagagaaatccatcc	240
Db	166517	ACCACAACCTTGCGCTTCAGCGCCAAACCCAGCCAGGGGCCAAGGGTGAGGAATGTCATCC	166458
QY	241	ggaagctctcatalggagcagcttgaggaaaccttgggcctcaattcccacatcttgagccgct	300
Db	166457	GGAGTCTCATATGGCCAGCTGGGAGAACCTTGGGGTCAATTTCCCACTCTTGAGAGCCGCT	166398
QY	301	atgacacagctgacacaccttccaccctacgcgtacatgcatgcccctgtgcatatggtgcataag	360
Db	166397	ATGACCACTGACACACTTTCACCTCCGCTACTGTGATGCCCCCTGTGCATATAGTGTCTAAGG	166338
QY	361	agcaaatgaggggaggagcagaagaaagaaagcccaactctcaagccttgaggcctgcgcc	420
Db	166337	AGCAAAATGGGGGGAGGCCAGGAGAAAGAACCCCACTTCTCAGGCGCTGGGGGCGTCCCC	166278
QY	421	actgtcctgtctcccaagctcccaagctgtctcctcaagcaagagacactlbggaaggtgggga	480
Db	166277	ACTGTCCGTGTTCCCAAGATCCGCCACTGTGTCTCAGCAACAAAGACACTGGAGGGGTGGGA	166218
QY	481	ggggagatctgacccctcaaacctgaccttccaccacaagagcccgaggcttgacctctcccgcc	540
Db	166217	GGGGATCTGAGACCTCAACCTGTCCCTTCACCCAAAGGCCCGGGCGTCAACCTCCGCCCGCC	166158
QY	541	ccctccctcagaggaatgctgacacatacgtacttcatctccgtgtggtgcatacgtactataa	600
Db	166157	CCCT--CCGCGAGAGGAGTGTGACCATCTACTTCACTCCGTGGTGTGCAATCGCTACTATA	166099
QY	601	ggtccccccttgggtgagagatggccttggaactcttatatggacaccacaagagcgaggggcgg	660

Db 166098	GGTCCCCCCCTGGTGAGAGAGTGGGCTGGGAATCTTTATGGGCAACCCAGAGGGCGGGGGCCG	166039
QY 661	aggagagtcctcccttgagagcctggtgcctctagaagcccaacgctctctcttgactctctgagtc	720
Db 166038	ACGGGAGTCTCTCTGGAGGCTCGTCCCTCCTAAGAGCCCACTTTCTGTGACTTCTGGAGTC	165979
QY 721	ctctgcattctcttaagttccagatctatagattgaacaactaagatctgaagctcgag	780
Db 165978	CTGTGCATGTCTCTAGGTCCAGATCTATAGATTGAACAACATTAAGATTGACACTGGAG	165919
QY 781	gggttgagttggggaaggaacccggaaaggggctctgtctgtgactgtgtgagctggaggtctlaag	840
Db 165918	GGGTGAGTGGGGAGAGGAACCCGGAAA--GGCGTTGGTGGATGGTGGCGGAGGCTTAAAG	165860
QY 841	gttggaagatagggaagtgagggaatgtctcttgaaatgaaagaaggaggaggaacaaatagaagctcg	900
Db 165859	GTGGAGAGATGGGAGTGGGAGTGTCTGTGAGTGAACAGGGGAGGGCAATTAAGAGACTTCG	165800
QY 901	gtgtccctgaacggaaggaaagctgtccctggagactgtcaaggtgaagcaggtgaccggtctccc	960
Db 165799	GGTGCTTACAGGAAGGAAGGAGTGCCTGGGAGTGAAGGTGAAGGAGGTGACCGGCTCCCC	165740
QY 961	tggcctgaactctggtcctctctctgcagaagatctcaactgtcaagagctgcttaagcctctggt	1020
Db 165739	TGGCTGACTCTGGCTCTTTCGTGCAAGATTCACCTGCAAGGCTGCTTCACACCTCTGGT	165680
QY 1021	ctccatctctctctaaatgaatggttggtgtctgtctgtcatctcaacctggagcctctctctcc	1080
Db 165679	CTGCATCTCTCCATATGATAGGTGGGTGTCTCTGCTCATTCACCTGAGCTCTCTCTCC	165520
QY 1081	acaagtcctctcccaagctcccaactcaagctctgaactcaactctcatctcaactaaggcgaca	1140
Db 165619	ACAGTCCCTTCTCCAGCTCCCACTAGCTCTTAACCTCACTTATCTTAAGCGCGACA	165560
QY 1141	cagacaaggagagccttggtgtccctgcacccctcttcttaaggagcctggagttgaggtgtgct	1200
Db 165559	CAGACAAGGAGACCTTGGTGGCTCTGCTCTCTTTTAAAGGGCTTGAGTGGAGTTGTCT	165500
QY 1201	ctccctaggctgtgcccgaggtcaactgtctccatctctctgcagcctgtcttgaagctgtcta	1260
Db 165499	CTCCCTTAGGCTCTCCCGAGGCTCACTGTGCTCCCATCTGTGACACCTGTTTGAAGCTGTCTA	165440
QY 1261	caacctctcaagaagaccggaatccatctgcgcctgtctgtcttcttgaacgggtgtcaaggaagct	1320
Db 165439	CACCTCATCAAGAACCGGATCATCTGCTGCTGCTTTTGAACCGGCTCAGGCAACCT	165380
QY 1321	actccacatctctacacacaacaaagcctgtcctcaagttctctgtcacatcttctgaagccttgg	1380
Db 165379	ACTCCACATCTCTACACACAAACGGCTGCTCAAGTTCGTGCAATCTTTGTAAGCTGTGG	165320
QY 1381	cccgaagtgggaagagggggagaccttgggcaggtgtatcaagaagggtctgaagatcttaag	1440
Db 165319	CCCAAGTGGGAAGAGGGGGAACCTTGGGAGGTGATAGAGGGCTCTAGAGAGCTTTCAG	165260
QY 1441	cccttagcagctgttggggaagagctctggagacccctcttgaagctgtctgtgaatctctatctcc	1500
Db 165259	CCCTAGCAGTCTCTGGGGAAGAGCTGGAGACCCCTCTTGAAGCTGCTGGATCCCTATATCTCC	165200
QY 1501	acctgtgtcccatctctaaacsaaggtctccctgtcgtcccgccgacctctctctctacacgac	1560
Db 165199	ACCTGTGCTCCCATCTTAACACAGGGTTCCTCTGCTGCCCCGGCCCTTCCTTAACCGAC	165140
QY 1561	tatccaagaatttgggcatacggcacatctccaggaacttggctgtgtgtctgtgaagacaacc	1620
Db 165139	TATCCAAATTTTGGGCATCTGGCACATTCCTCGAGACTTGGCTGTGTGTGCGAGACAGCAC	165080
QY 1621	catctgtgactgtcacctgtgacatctcttgtgaacgggcgtgtgtctgtcaactgtccctgtgtcaa	1680
Db 165079	CATCTGTACCTGACACTGGACATCTTTGTGTGACCGGCGTGTCTGTGCATCTGCTGTGCTCAA	165020
QY 1681	cgaatgtgtgtacccaacccccagaatgaagagctgtggtgtg 1721	
Db 165019	CGAATGTGTGTACCCACCCCGAGATGAGAGGCTGTGGGCTGG 164979	

RESULT 5
AF214521 5888 bp DNA linear MAM 03-JUN-2000
LOCUS Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
DEFINITION AF214521
ACCESSION AF214521 GI:8215685
VERSION AF214521.1 GI:8215685
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 5888)
AUTHORS Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
Kogel-Galliard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
Kogel-Galliard,C., Paul,S., Iannucelli,N., Rask,L., Rome,H.,
Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
and Andersson,L.
A mutation in PRKAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)
JOURNAL MEDLINE
PUBMED 20280150
REFERENCE 10818001
AUTHORS 2 (bases 1 to 5888)
Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
Kogel-Galliard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
Sweden
FEATURES
source Location/Qualifiers
1..5888
/organism="Sus scrofa"
/db_xref="taxon:9823"
/chromosome="15"
/map="15q"
join(1..154,515..918,1809..1890,2349..2407,2509..2554,
2771..2825,3027..3153,3286..3451,4578..4615,4791..4937,
5294..55410)
/gene="PRKAG3"
/product="AMPK gamma subunit"
1..>5410
/gene="PRKAG3"
join(1..154,515..918,1809..1890,2349..2407,2509..2554,
2771..2825,3027..3153,3286..3451,4578..4615,4791..4937,
5294..5410)
/gene="PRKAG3"
/codon_start=1
/product="AMPK gamma subunit"
/protein_id="AA073989.1"
/db_xref="GI:8215686"
/translation="MSFLEQGSRSMPRAVTTSSERSHGDNKASHMTROEDVEEG
GPPREGPOSPVAESTGOEATPEPKAPLAOAALAEVNDPPTREDRLLPSCAASAS
DSNTDLDLGIEFSASASGDELGLVEKRPAPCPREVLRLRLGDDDLQVPGAOVYM
HEMOBHTQDAMATSSKLVITDPTMLEIKAPFALYANVRAPLYDSKKOSFVCHLT
TDFILVHRYRSPVQVQYEIEBKETETWRELYDGCKPLVTSFNDLSLEAVYALI
KNRIHRLPVLDPVSGAVLHILHNRKLFHIFGLLPSPFLVKTIDDLIGTFRL
AVULETARILTALDLIFDVRVSALEPVNVEQVGLVSRFVLIHLAQQTYNHLDMY
GEALRORTLCLEGLVLSQPHNETLEVIDRIIVREQVNRVLVDEQHLGLVSLSDILO
ALVSPACIDALIGA"
BASE COUNT 1274 a 1634 c 1638 g 1339 t 3 others
ORIGIN

Query Match 35.6%; Score 612.6; DB 4; Length 5888;
Best Local Similarity 66.3%; Pred. No. 1e-137;
Matches 1155; Conservative 0; Mismatches 484; Indels 102; Gaps 15;

QY 5 gccctcagatcaagaagcctcttgcctctggtgagccacggtgctgcggcagccct 64
DB 1800 GGCCTCAGATCAAGAAGCCTTCTTGGCTGTGGCCACGCGCTCCGACGCGACCT 1859

QY 65 ctaatgagcagcaagaagcagccttctggtgaggaagcctgaggaagtgaaaggag 124
DB 1860 TTGTGGACAGCAAGAAGCAGACTTCGTGGCTGAGAGCGCTGGGAGG-CAGAGGTG 1918
QY 125 atgaggaagtgaggaaggaatcttgcacgtgtgtcttcggggcttgatctctgatalacca 184
DB 1919 GTGGGGAAGGGAATAGGGGACCTTGTGGGTGATTCTTAAGGCCACAGCTCTACACACCA 1978
QY 185 caagcttgctcacaagccagccagggcgaggaagtgaggaagatccatccaggag 244
DB 1979 CAGGCTTCAACCAACAGCAGGCGCTGGCCGTGAG--AGGGGGGACCATTTTACCCCGGCT 2036
QY 245 tctgcatgagcagcctgaggaagccttgggtgctcaattcccatctgagagcgtatga 304
DB 2037 TCGTGGTGGCCAGCTGGGAGA-----TCTCAACTGTAGAGACAGCTGTGA 2080
QY 305 caagctggaaccttccacctccgctacatgcagtcgacctgtgca-----tagtgctaa 358
DB 2081 CCAGCTGACCCCTCCAGCTCTACTACCCCAAGGTCCTTCGACGTCTAGTAAGAAAG 2140
QY 359 ggaagcaaatgaggaagcaggaagaaagcccaacttccagcctgagggcgtgcc 418
DB 2141 AGGACAGCGCGAGGAAGGAAGTACAGAAATAGAAAGCAGGCGCAGGAAGGAAGAAATG 2200
QY 419 ccactgtctgtcccaagatcccaactgtgtctcagaacaagaaactgaggaagtgag 478
DB 2201 ACAGGGGAAGCAATAAGAGGACACCCCATTTGTACGGCAGGAGGGGCTCCCTCCCT 2260
QY 479 gaggagatctgacctcaaccgtcc-----ttcacccaaagccccggcct 525
DB 2261 TCCCTCTTTTGGCCACCCTCAGTAAAGATGTGGCAGGTCGGGGGAGGGCCGCGCT 2320
QY 526 gacctcctcccccctc-cctctgaggaatgctgacaatcatcgaatctatcctgtg 584
DB 2321 GACCCCAATGCTCCTCCCTGCCCCACAGAGGATGCTGACCATACAGACTTCACTTGGTG 2380
QY 585 ctgcatcgctactacaagtcceccctgtgtgaggaatgagccttgaaatctlatggcaacc 644
DB 2381 CTGCACCGCTATTACAGGCTCCCTCGTGAAGAGTGTGGTGGGCTCT--GGAACACCC 2439
QY 645 aagaagcagggcgaggaagggagctcctctgagcctgtgtgcccagaagccacgtct 704
DB 2440 ATCTGGGCTGGGGTGGAAGGACTTC-----AGGGGACCTTCG 2476
QY 705 tctgactctgagtcctgtcgtatgtctctcagtcacagatctatagatltgaacaacata 764
DB 2477 CCTGACTTTGGAGTTCGCTGCTGCTCTCTTTAGCTCCAGTACAGAAATTGAAGAACATA 2536
QY 765 agattggaacctgaggaaggtgagtgaggaagaaacccgaaaggcctgtgtgt----- 819
DB 2537 AGATTGAGACCTGGAGGGGTGAGCAGCGAGCGGCGGCGGAGGGGCTGAGGGTGTGTG 2596
QY 820 -----gagtgaggccaaaggtctaaagtgaggaatgaggaagtgagtgagtgagga 874
DB 2597 GGTGAGGATGGGGCCAAAGGACTCTACGGAGACATGCGCAGTGGAGGTTCTCTGAGGAA 2656
QY 875 acaaggaaggaacaatlaagagcctctggtgtcctgaacggaaggaagcctgctggagcty 933
DB 2657 GCGGAGGAGGGGTATTCGGAGACCCACGGGATCTAAGGGAGGAGACAGTGTGGGGGTGG 2716
QY 934 -caagtgaggaagtgagcagcgtctccctctgagcctgagcctgtgtctcttcgagaagatct 992
DB 2717 CCAGCTGAGGGCGGGGTGTGGCCCTTGTGTGCTGATCTTGCGCTTTTCTCCAGAGATCT 2776
QY 993 acctgcaaggtctcctaagcctctgtctcatctctccatctcctaataatagatgagtgctc 1052
DB 2777 ACCTTCAAGGCTGCTTCAAGCCTCTGTCTCATCTTCTCCCAATGACAGGTGAGCTTCC 2836
QY 1053 tgcctaatcacttgagcctctctctccacaagctccctctcccaagtlccacatcagctgt 1112
DB 2837 CAGCGCCCACTGAGGCTCTCTGCGCCGACAGACCCCTTCTCCAGCTCATTCGGTTCTA 2896
QY 1113 aactcaactcttcaatcctcagaagcgagcaacaagaaggaagccttggtgcctgcctcct 1172

TITLE	FEATURES	BASE COUNT
<p> Variant of the human amp-activated protein kinase gamma 3 subunit Patent: WO 0177305-A 18-OCT-2001; Araxis AB (SE) </p>	<p> source 1. 1647 Location/Qualifiers /organism="Homo sapiens" /db_xref="taxon:9606" 20. 1489 /note="unamed protein product" /codon_start=1 /protein_id="CAD10589.1" /db_xref="GI:16608834" /translation="MEGLEHALRRTPSMWSLGCSEHOEMSLFEQNSSMSPSPAVTSSSEIRKGRRAKRALRRTRRQKSVSEGEPPGCGEGRSRPAAESTLEATFPKTTPLAOLAPDPAVGTPPGWDCILPSDCAASAAGSSITDVELTEPPATEAMCEEGLEERPAFLSPDAPFPKLGMDLELRRKPGAQIYMRQMEHTCYDAMATSKLVIPTMLEIKKAPFALVAGVNAAPLAMDSSKOSFQOMLTITDFILVLHYNSPLVQIYEITDHRKLETMRLEYLVGGEFKPLVASSPNDSSLFEAVYTLIKRNIHLPLVDPGSGVLAHILHKLKTLFIYGLLRPSPLRYRTIODLGTGFTRDLVALETPALITLADIFVDRYSNALPVYNGCOVVGSLSPRPDIYHLAAQYTNHNDMSVGCALRRTLCESVLSQPHESLGEVIDRIAREQVHRILVLDTEHLLGVSSLDLIDQALVLSAPGIDALGA" </p>	<p> 346 a 502 c 462 g 337 t </p>

Query Match	9.88;	Score 168.2;	DB 6;	Length 1647;
Best Local Similarity	95.68;	Pred. No. 4.2e-30;		
Matches 173;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
QY 1510	ccatcctaaccaagggtctccctgctgcgccggccctctcctctacgcgactatccaaga	1569		
Db 1009	CCTGCACATCTTTGGTTCCCTGCTGCCGCCGCCCTTCCTCTACCGCCACTATCCAAAGA	1068		
QY 1570	cttgggcattcggcacatctccgaagacttgcgcgtgtgtctctgtagagacagccatctgac	1629		
Db 1069	TTTGGCGATCCGCCACATTCACAGACTTGGCGTGTGGTGTGAGACAGCACCCATCTCTAC	1128		
QY 1630	tgcactgacacatctcttggagccggcggtgtctgcgcaactgctctgtgtcacaagaattg	1689		
Db 1129	TGCACCTGCACATCTTTGTGTGAGACCGGCCGTGTCTGTGCATCGCTGTGTGTCACAGAAATGTGG	1188		
QY 1690	t 1690			
Db 1189	t 1189			
RESULT 8				
AX099776				
LOCUS	AX099776	2109 bp	DNA	Linear
DEFINITION	Sequence 3 from Patent WO0120003.			
ACCESSION	AX099776			
VERSION	AX099776.1	GI:13538810		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2109)			
AUTHORS	Andersson, L., Looft, C., Kalm, E., Milan, D., Rodic, A.,			
	Rogel-Galliard, C., Iannuccelli, N., Gellin, J., Le Roy, P. and			
	Charbon, P.			
TITLE	Varants of the gamma chain of ampk, dna sequences encoding the			
	same, and uses thereof			
JOURNAL	Patent: WO 0120003-A 3 22-MAR-2001;			
	INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;			
	Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)			
FEATURES	Location/Qualifiers			
Source	1..2109			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	472..1389			
CDS	/note="unnamed protein product"			
	/codon_start=1			

[illegible][illegible]

BASE COUNT	460 a	622 c	562 g	471 t
ORIGIN	ALVSPAGIDALTA*			
Query Match	9.88;	Score 168.2;	DB 9;	Length 2115;
Best Local Similarity	95.68;	Pred. No. 4e-30;		
Matches 173;	Conservative	0;	Mismatches 8;	Indels 0;
			Gaps	0;
Qy 1510	ccatcctaaccagggtccctgctgcgcgcgcgcctctctctcaacgcgactatcaaga 1569			
Db 915	CTGTCAATCTTTGGTTCCTGCTGCTGCCGCCCTCTCTCTTCAACGCATATCAAGA 974			
Qy 1570	cttggcgcgcgcacatccgcagactgctgctgctgctgagacagaccactctgac 1629			
Db 975	TTTGGCATCGGCACATTCGAGACTTGGGCTGTGTCTGTCTGAGACACACCATCTGAC 1034			
Qy 1630	tgcactgacatctctgtgagacggctgtgtctgcactgctgctgctcaagaatgtg 1689			
Db 1035	TGCACTGCATCTTTGTGTGACCGCGCTGTGTGTGACCTGCTGTGTCACCAAGATGTG 1094			
Qy 1690	t 1690			
Db 1095	t 1095			
RESULT 11	HSA249977 2290 bp mRNA linear PRI 07-Apr-2000			
LOCUS HSA249977				
DEFINITION	Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit (AMP gamma 3 gene).			
ACCESSION	AJ249977			
VERSION	AJ249977.1 GI:6688200			
KEYWORDS	AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit. human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2290)			
AUTHORS	Cheung,P.C., Salt,I.P., Davies,S.P., Hardle,D.G. and Carling,D.			
TITLE	Characterization of AMP-activated protein kinase gamma-subunit isoforms and their role in AMP binding			
JOURNAL	Biochem. J. 346 Pt 3, 659-669 (2000)			
MEDLINE	2016049			
REFERENCE	2 (bases 1 to 2290)			
AUTHORS	Carling,D.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC Clinical Sciences Centre, Hammersmith Hospital, DuCane Road, London, W12 0NN, UNITED KINGDOM			
FEATURES	Location/Qualifiers			
source	1..2290			
gene	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	22..1500			
	/gene="AMPK gamma 3"			
	22..1500			
	/gene="AMPK gamma 3"			
	/function="AMP-activated protein kinase regulatory subunit"			
	/codon_start=1			
	/evidence="experimental"			
	/product="AMP-activated protein kinase gamma 3 subunit"			
	/protein_id="CAB65117.1"			
	/db_xref="GI:6688201"			
	/translation="MEPCIEHLARTRPSSSLGSGSEHQBMSLEEDNSSSPSPAVTS			
	SSEIRIGRRRAKALRWTROKSVEESEEPQGGCPSPSPAESTGLEATFPKTTPLAQA			
	DPAGQPTPTGDCIDLPSCDTASAGSSTDVDLAIEFATMECELEGLLEERPALQ			
	LSPQAPFKLIGWDELKRPGAQIYVRFLTEHTCYDAMATSSKLVIFDTNLTKKAFFA			

Query Match	8.6%	Score 148;	DB 6;	Length 1867;
Best Local Similarity	88.9%	Pred. NO. 3.2e-25;		
Matches 160; Conservative	0;	Mismatches 20;	Indels 0;	Gaps 0;

Query Match 8.68; Score 148; DB 4; Length 1873;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 12:32:10 ; Search time 719.93 Seconds
(without alignments)
4106.686 Million cell updates/sec

Title: us-09-826-581-3

Perfect score: 1722

Sequence: 1 cctggccctcagatcaaga.....gatgagagctcggtcgtga 1722

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapexl 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1722	100.0	1722	22	AAH43683
2	351	20.4	378	22	ABAA44706
3	351	20.4	378	22	ABA45162
4	351	20.4	378	22	ABA24907
5	351	20.4	378	22	AAK03423
6	351	20.4	378	22	AAK28874
7	351	20.4	378	22	AAI13465
8	351	20.4	378	22	AAI14821
9	351	20.4	378	22	AAI03344

10	196	11.4	547	22	ABA08485	Human AMP-activate
11	168.2	9.8	1647	22	AAH43685	PRKAG3 CDNA. Homo
12	168.2	9.8	2109	22	AAAD03296	Human AMPK gamma s
13	168.2	9.8	2115	22	AAAD03320	Human AMPK gamma s
14	148	8.6	1867	22	AAAD03295	Pig AMPK gamma sub
15	148	8.6	1873	22	AAAD03319	Pig AMPK gamma sub
16	148	8.6	2022	22	AAAD03321	Sus scrofa PRKAG3
17	74.2	4.3	16525	22	AAK73303	Human immune/thema
18	70.4	4.1	602	22	AAH35203	Human colon cancer
19	70.4	4.1	1691	21	AAAC98774	Human pancreatic c
20	68.8	4.0	1576	18	AAT85927	Mammalian AMPK-gam
21	68	3.9	39651	23	ABLI18856	Drosophila melanog
22	65.8	3.8	3261	23	ABLI18857	Human breast cell
23	65	3.8	92	22	ABA49850	Human foetal liver
24	65	3.8	92	22	ABA67769	Probe #13292 for g
25	65	3.8	92	22	ABA4826	Human brain expres
26	65	3.8	92	22	AAK16181	Human bone marrow
27	65	3.8	92	22	AAK41922	Probe #12625 for g
28	65	3.8	92	22	AAI22692	Probe #16674 used
29	65	3.8	92	22	AAI47988	Probe #8345 used t
30	65	3.8	92	22	AAI08354	Disease associated
31	63.8	3.7	1435	20	AAK06882	DNA encoding novel
32	63.8	3.7	1467	23	AAK84265	Human CDNA sequenc
33	63.8	3.7	2223	22	AAH14839	Human secreted pro
34	62.6	3.6	350	21	AAAC01661	Human CDNA clone (
35	57.2	3.3	735	22	AAH07561	DNA encoding novel
36	49.8	2.9	2303	23	AAK84267	Mouse ischaemic co
37	49.4	2.9	151	24	AB199496	Novel human polynu
38	44	2.6	375	22	AAK6844	Mouse cancer assoc
39	42.6	2.5	985	22	AAAD08652	Human immune/thema
40	42	2.4	61710	22	AAK3782	Human immune/thema
41	41.6	2.4	37664	22	AAK3782	Human secreted pro
42	40.8	2.4	1512	21	AAZ98024	Human secreted pro
43	40.8	2.4	1512	21	AAI1637	Arabidopsis thalia
44	40.6	2.4	766	21	AAK4215	Human nervous syst
45	40.6	2.4	8017	22	ABAI9015	

ALIGNMENTS

RESULT 1	AAH43683	standard; DNA; 1722 BP.
XX	AAH43683;	
XX	21-JAN-2002	(first entry)
XX	PRKAG3 intron 4 -	intron 10.
XX	Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;	
KW	metabolic disease; diabetes; obesity; substitution; ds.	
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
XX	intron	1..13
FT	intron	/tag- a "intron 4"
FT	exon	/number- "3' portion of intron 4"
FT	exon	14..95
FT	exon	/tag- b
FT	intron	/number- "Exon 5"
FT	intron	96..552
FT	intron	/tag- c
FT	exon	/number- "intron 5"
FT	exon	553..611
FT	exon	/tag- d
FT	exon	/number- "Exon 6"
FT	intron	612..736
FT	intron	/tag- e
FT	intron	/number- "intron 6"

```
FT      exon      737..782
FT      /tag= f
FT      /number= "Exon 7"
FT      Intron    783..986
FT      /tag= g
FT      /number= "Intron 7"
FT      exon      987..1041
FT      /tag= h
FT      /number= "Exon 8"
FT      Intron    1042..1242
FT      /tag= i
FT      /number= "Intron 8"
FT      exon      1243..1369
FT      /tag= j
FT      /number= "Exon 9"
FT      Intron    1370..1522
FT      /tag= k
FT      /number= "Intron 9"
FT      exon      1523..1688
FT      /tag= l
FT      /number= "Exon 10"
FT      Intron    1689..1722
FT      /tag= m
FT      /number= "Intron 10"
FT      /note= "5' portion of Intron 10"

PN      WO200177305-A2.
PD      18-OCT-2001.
XX      06-APR-2001; 2001WO-SE00765.
XX      07-APR-2000; 2000US-195665P.
PA      (AREX-) AREXIS AB.
PI      Andersson L, Luthman H, Marklund S;
XX      WPI: 2001-657170/75.
XX      New variants of human AMP-activated protein kinase gamma3 subunit
PT      associated with a metabolic disease e.g. diabetes or obesity and method
PT      for determining a risk estimate of diseases in subject by detecting the
XX      variant.
XX      Example 1; Fig 3; 25pp; English.
PS      The sequences given in AAH43681-84 represents genomic fragments
CC      encoding the human AMP-activated protein kinase gamma 3 subunit
CC      (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC      is useful in determining a risk estimate of a metabolic disease,
CC      such as diabetes or obesity, in a subject. The variation may occur
CC      in exons 3, 4 or 10. In exon 3 variation may be a substitution of
CC      a G for a C at nucleotide 320, resulting in the amino acid
CC      substitution P71A: in exon 4 variation may be a substitution of a
CC      T for a C at nucleotide 550, and in exon 10 variation may be a
CC      substitution of a T for a C at nucleotide 1037, resulting in the
CC      amino acid substitution R340W. There may also be nucleotide variation
CC      in Intron 6.
XX      Sequence 1722 BP: 321 A; 504 C; 534 G; 363 T; 0 other;

Query Match      100.0%; Score 1722; DB 22; Length 1722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 cctggccccacagataagaagcctcttctcgtggtgagccacggtgtgagggcagc 60
Db      1 cctggccccacagataagaagcctcttctcgtggtgagccacggtgtgagggcagc 60
QY      61 cctcctatgagacagcaagaacagagcttctggttgagaggaagctcggggaagtgaag 120
Db      1141 cctcctatgagacagcaagaacagagcttctggttgagaggaagctcggggaagtgaag 1200
1141 cagacaaaggagagccttctgagcctcctccttcttlaaggagcctgagatgaggtgct 1200
```

```
Db      61 cctcctatgagacagcaagaacagagcttctggttgagaggaagctcggggaagtgaag 120
QY      121 gggagatgaggaaggtgagggggagatctctgacggtgtctcggggctgatatctgatat 180
Db      121 gggagatgaggaaggtgagggggagatctctgacggtgtctcggggctgatatctgatat 180
QY      181 acccaagacttgctcacaagccagccagggggccagagggtgaggaagaatccatcc 240
Db      181 acccaagacttgctcacaagccagccagggggccagagggtgaggaagaatccatcc 240
QY      241 ggaagctcgaatgagcagctcggagagaccctcggggtccaaatctccatctgtgagacgct 300
Db      241 ggaagctcgaatgagcagctcggagagaccctcggggtccaaatctccatctgtgagacgct 300
QY      301 atgacacagctgacacaccttccacctcgcgtacgtacgtacgtacgtacgtacgtacgtacgt 360
Db      301 atgacacagctgacacaccttccacctcgcgtacgtacgtacgtacgtacgtacgtacgtacgt 360
QY      361 agcaaatgggggagagcaggaagagaagaagaccacacctcctcagagcctgggggctgcgcc 420
Db      361 agcaaatgggggagagcaggaagagaagaagaccacacctcctcagagcctgggggctgcgcc 420
QY      421 actgtctctgtcccaagctcccaagctctgtctcagacagaagacatgagcaggttgagga 480
Db      421 actgtctctgtcccaagctcccaagctctgtctcagacagaagacatgagcaggttgagga 480
QY      481 gggagatcgaacctcaacctgtcctccacaaaggccggggtctgacctctcccgcc 540
Db      481 gggagatcgaacctcaacctgtcctccacaaaggccggggtctgacctctcccgcc 540
QY      541 cctccctcgaaggaatgctgacacatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 600
Db      541 cctccctcgaaggaatgctgacacatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 600
QY      601 ggtcccccctggtgagaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 660
Db      601 ggtcccccctggtgagaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 660
QY      661 agggagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 720
Db      661 agggagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 720
QY      721 ctgtcgaatgtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 780
Db      721 ctgtcgaatgtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 780
QY      781 ggtgtagtgaggagaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
Db      781 ggtgtagtgaggagaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
QY      841 gtggaagatgagcagtgaggatgtcctcgtgagtgaaacagggagggaggaacaataagagctc 900
Db      841 gtggaagatgagcagtgaggatgtcctcgtgagtgaaacagggagggaggaacaataagagctc 900
QY      901 ggtgctcgaaggaaggaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 960
Db      901 ggtgctcgaaggaaggaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 960
QY      961 tggcctgactctggtctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1020
Db      961 tggcctgactctggtctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1020
QY      1021 ctccatctcctccaaatgatatggtgtgtctcgtcctcctcctcctcctcctcctcctcctcct 1080
Db      1021 ctccatctcctccaaatgatatggtgtgtctcgtcctcctcctcctcctcctcctcctcctcct 1080
QY      1081 acagtcctcctcccaagctcccaagctcccaagctcccaagctcccaagctcccaagctcccaag 1140
Db      1081 acagtcctcctcccaagctcccaagctcccaagctcccaagctcccaagctcccaagctcccaag 1140
QY      1141 cagacaaaggagagccttctgagcctcctcctcctcctcctcctcctcctcctcctcctcctc 1200
Db      1141 cagacaaaggagagccttctgagcctcctcctcctcctcctcctcctcctcctcctcctcctc 1200
```

Oy	1201	ctctttagtggtcccccgaagcgtacagtcctccatctctgcagaacctgtttgaagtgtcta	1260
Dh	1201	cctccctaaagctcgccccgaaggcttaactgctcccatctctgcagaacctgtttgaagtgtcta	1260
Oy	1261	caaccctacaagaataaacccggatccattctgctctgtcttctctgaaccgggtgatagaacaagt	1320
.b	1261	caaccctcatcaagaataaacccggatccattctgctctgtcttctctgaaccgggtgatagaacaagt	1320
Oy	1321	actccacaatccctcacacaacaacagcctgtctcaaatgtctctgcacatcttgttaaagcttggg	1380
Lb	1321	actccacaatccctcacacaacaacagcctgtctcaaatgtctctgcacatcttgttaaagcttggg	1380
Oy	1381	cccgagtgaggagaaagaggaggagcccgaggcaagtgatcaagaagagcgctgaagaagcttcag	1440
Dh	1381	cccgagtgaggagaaagaggaggagcccgaggcaagtgatcaagaagagcgctgaagaagcttcag	1440
Oy	1441	ccctcagcagctcgctggagggaagagctctggagaacctctctgaagctgcagctgaatctctcc	1500
Dh	1441	ccctcagcagctcgctggagggaagagctctggagaacctctctgaagctgcagctgaatctctcc	1500
Oy	1501	aactgtgtcccatctccaaccaagggtctcctgtctgccccggagccctctctctctacacgac	1560
Dh	1501	aactgtgtcccatctccaaccaagggtctcctgtctgccccggagccctctctctctacacgac	1560
Oy	1561	tatccaagaatttgagcatcgcgcacatctccgagaacttgctgtgtgtctgtagagacagcacc	1620
Dh	1561	tatccaagaatttgagcatcgcgcacatctccgagaacttgctgtgtgtctgtagagacagcacc	1620
Oy	1621	catctcgtactgcacatcgacatcttltgtgagccggcgtgtgtctgcactgtcctgtgtcataa	1680
Dh	1621	catctcgtactgcacatcgacatcttltgtgagccggcgtgtgtctgcactgtcctgtgtcataa	1680
Oy	1681	cgaaatgtgtgtacccaccaccccaagatagaagagcttgggccttga 1722	
Dh	1681	cgaaatgtgtgtacccaccaccccaagatagaagagcttgggccttga 1722	
RESULT 2			
ABAA44706/c			
ID	ABAA44706 standard; DNM: 378 BP.		
XX	ABAA44706;		
AC			
XX			
DT	01-FEB-2002 (first entry)		
XX			
DE	human breast cell single exon nucleic acid probe #3401.		
XX			
KW	Human; microarray: single exon probe; gene expression; breast; disease: cancer; ss.		
XX			
OS	Homo sapiens.		
PN	WO200157271-A2.		
PX			
PD	09-AUG-2001.		
PX			
PF	30-JAN-2001; 2001MO-USO0662.		
PX			
PR	04-FEB-2000; 2000US-0180312.		
PR	26-MAY-2000; 2000US-0207456.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0633366.		
PR	21-SEP-2000; 2000US-0234687.		
PR	27-SEP-2000; 2000US-0236359.		
PR	04-OCT-2000; 2000GB-0024263.		
PX			
PA	(MOL-E) MOLECULAR DYNAMICS INC.		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX			
DR	WPI; 2001-496933/54.		
XX			

PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS
XX Claim 1; SEQ ID NO 3401; 327bp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;

Query Match	20.48; Score 351; DB 22; Length 378;
-------------	--------------------------------------

Matches	362;	Conservative	0;	Mismatches	0;	Indels	1;	Gaps	1;
---------	------	--------------	----	------------	----	--------	----	------	----

QY 1 ccctgccccctcagatcaagaagccctcttctgtctgtgcacaagtgtgcggcagc 60
|||||
Db 362 CCTGCCCTCAGATCAAGAAGCCCTTTGCTGTGGTGCACAACGCTGTGGCGCAGC 303

Qy 61 cccctatgagacagcaagaagcagaacttctgtgtgaggaagagcgtctgggaggtgaag 120
 |||||
 Db 302 CCCCTCTATGGCAGCAGCAAGAAGCAGACCTTTGTGGTGTAGAGAGAGCGCTGGGGAGAGTAA 240

Oy 121 ggaaatactgaagatctgaaggaggagatactctgaacatctcgtat 180
 |||||
 Db 242 GGAATGGAGAGGTGAGGGGAGATCTTGTACGGTTCTTGGGCTCATCTCATAT 183

QY 181 acccaagcttggtcttaagcacaagccagcagggcagggcgagggggaagtcacatcc 240
 DB 182 ACACACAAGCTTGCGCTTCAGGCAAGCCACGCCAGGCGCCAGGGCCAGGGTGGAGGAAGTCATCC 123

[illegible]

0Y 361 aac 363
 Dd 62 ATGACACAGCTGACACCTTTTCACTCCCTACTGCATGGCCCTGTG-CATAGTGCTAGGG 4

Db	3	AGC	1
----	---	-----	---

ABA55162/c	ABA55162 standard; DNA; 378 BP.
ID	
XX	
AC	ABA55162.

XX	
DT	01-FEB-2002 (first entry)
XX	
DE	Human foetal liver single exon nucleic acid probe #3467.

Human; foetal liver; gene expression; single exon nucleic acid probe; ss

OS	Homo sapiens.
XX	
PN	WO200157217-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00669.
XX	
RP	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
DR	WPI; 2001-483447/52.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human fetal liver -
XX	
PS	Claim 1; SEQ ID NO 3467; 639pp + sequence listing; English.
XX	
CC	The invention relates to a single exon nucleic acid probe for
CC	measuring human gene expression in a sample derived from human foetal
CC	liver. The single exon nucleic acid probes may be used for predicting,
CC	measuring and displaying gene expression in samples derived from human
CC	fetal liver. The present sequence is a single exon nucleic acid
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at fcp.wipo.int/pub/published_pct_sequences.
XX	
SO	Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;

[illegible]

XX	RESULT	4
XX	ABA24907/c	
ID	ABA24907	standard; DNA; 378 BP.
XX		
AC	ABA24907;	
XX		
DT	23-JAN-2002	(first entry)
XX		
DE	Probe #3373	for gene expression analysis in human heart cell sample.
XX		
KW	Human; gene expression; heart; microarray; vascular system; probe;	
KW	cardiovascular disease; hypertension; cardiac arrhythmia;	
KW	congenital heart disease; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200157274-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001MO-US00666.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236559.	
PR	04-OCT-2000; 2000GB-0024263.	
XX		
PA	(MOLE-)	MOLECULAR DYNAMICS INC.
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-48899/53.	
XX		
PT	Single exon nucleic acid probes for analyzing gene expression in human	
XX	hearts -	

PS Claim 1 SEQ ID NO 3373: 530pp: English.

xx CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://wipo.int/publ/published_pcl_sequences](http://wipo.int/publ/published_pcl_sequences).

xx CC
xx CC
xx CC

Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other:

Query Match 20.4%; Score 351; DB 22; Length 378;
Best Local Similarity 99.7%; Pred. No. 6,96-83;
Matches 362; Conservative 0; Mismatches 0; Indels 1; Gaps 1

1 ccctgcctccatcaataaagagctcttcctgctcctggtgccaaggtgtgcygagc 60
362 cctggccctctcagatccatcaagagccctttcttgcctctgctgccaagctgctggcagc 303

61 ccctctatggagcaagcaagaagcagagctctgtgtgtatgaggaagagcctgggagtgtaag 120
121 ccccttatgtggagcagcagcaagcaagcagcttgtgtggtagagagagctgggagtgtaag 243

242 ggagatggagcagcagctgtagggagagctcttgaagagtggtcttggtggcttgatcctgatat 180
183 ggagatggagcagcagctgtagggagagctcttgaagagtggtcttggtggcttgatcctgatat 183

QY 181 accacaagcttggtcttcaggtccagccagcaggggtggaagaaagtcaccc 240
DB 182 ACCACCAAGCTTGCTTCAGGCCAAGCCCAAGCCAGGGGTGAGGAAAGTCCATCC 123
QY 241 ggaagctgcagtcagcagcctgagagaccctgagggctcaatttcccatctgtgagccgct 300
DB 122 GGAGTCTGCATGGCCAGCTGGAGACCTTGCGCTCAATTTCCCATCTGTGGAGCCGCT 63
QY 301 atgaccagctgacaccttccacctccgctactgcatgagccctgtgcatagtgctcaggg 360
DB 62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGTCATGCGCCCTGTG-CATAGGTGCTAGGG 4
QY 361 agc 363
DB 3 Agc 1
RESULT 5
AAK03423/C
ID AAK03423 standard; DNA: 378 BP.
XX
AC AAK03423;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 3414.
XX
KW Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PJ 09-AUG-2001.
XX
PI 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 3414; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SO Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;

Query Match 20.4%; Score 351; DB 22; Length 378;
Best Local Similarity 99.7%; Pred. No. 6,9e-83;
Matches 362; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ccttgccccctcagatcaagaagccttcttctgtgtggtccaaaggtgtcgggagc 60
DB 362 CTTGGCCCCCTCGAGATCAAAAGAGCCCTTCTTGTCTGTGGCCAAAGGTGTCCGGCAGC 303
QY 61 cccctcatalggagcagaagaagcagagcttctgagtgagagagagctgagagtgag 120
DB 302 CCTCTATGGAGACAGCAAGACAGCTTTGTGGCTGAGAGAGAGCTGGCAGGTGAAG 243
QY 121 ggaagctgagagagtgagagggagagatctgtacggttctctgagggctgatalcttgat 180
DB 242 GGAGATGAGAGAGATGAGGGGAGATCTGTACGGTTGTTCTGGGCTCATCTGATAT 183
QY 181 accacaagcttggtcttcaggtccagccagcaggggtggaagaaagtcaccc 240
DB 182 ACCACCAAGCTTGCTTCAGGCCAAGCCCAAGCCAGGGGTGAGGAAAGTCCATCC 123
QY 241 ggaagctgcagtcagcagcctgagagaccctgagggctcaatttcccatctgtgagccgct 300
DB 122 GGAGTCTGCATGGCCAGCTGGAGACCTTGCGCTCAATTTCCCATCTGTGGAGCCGCT 63
QY 301 atgaccagctgacaccttccacctccgctactgcatgagccctgtgcatagtgctcaggg 360
DB 62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGTCATGCGCCCTGTG-CATAGGTGCTAGGG 4
QY 361 agc 363
DB 3 Agc 1
RESULT 6
AAK28874/C
ID AAK28874 standard; DNA: 378 BP.
XX
AC AAK28874;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 3431.
XX
KW Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PJ 09-AUG-2001.
XX
PI 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 3431; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers

XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488897/53.
XX
PI Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 3507; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;

Query Match	20.4%	Score 351	DB 22	Length 378
Best Local Similarity	99.7%	Pred. No. 6	9e-83	
Matches 362	Conservative 0	Mismatches 0	Indels 1	Gaps
QY 1	cctggccctcaga tcaagaagcctctctctgtgctgtgagccaacggtgtgcggcagc	60		
Db 362	CCTGGCCCTCAGATCAAGAAAGGCGCTTCTTGCTCTGTGCGCAACGGTGTGGGCGAG	303		
QY 61	ccctctctgggaacgcaagaagcgaagccttctgtgtgtgaagaagccttgggaagtgaa	120		
Db 302	CCCTCTATGGACAGCAGAGAGCGACGACCTTGTGGTGAGCAGACGCTGGGAGGTGAAG	243		
QY 121	ggagatlggaagagtggaagggggaatcttgaagttctgtctgtgggctgtatctctgatal	180		
Db 242	GGAATGTGAGAGAGGTGAGGGGGAATCTTGTACGGTTGTTCTGGGGCTGATCTGTATAT	183		
QY 181	accgaagccttggcttcaaggccaagcccaacgaaggcggaagtggaagaagatccatcc	240		
Db 182	ACCCACAGCTTGCTGCTTAGGCGACAGCCACCGGCGCAGGGGTGAAGAAATTCATCC	123		
QY 241	ggaagctcgaatcgtgcaagcctgtggaagcccttgagggtcgaattcccaatcgtggaagccgt	300		
Db 122	GGAATCTGCATGGCGACACTGGAGACCCGTGGGGCTCAATTTCCCATMCTGTGAGCGCT	63		
QY 301	atgacgaagcctgaaaccttcaaccctcgcgtacgtacatgagccctgtgcacaagtgtcagaag	360		
Db 62	ATGACCAAGCTGACACCTTTCACCTCCGCTACTGCATGCGCCCTGTG-CATAGGTGTTAGGG	4		
QY 361	agc 363			
Db 3	AGC 1			

XX	inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX	
OS	Homo sapiens.
XX	
PN	MO200157270-A2.
XX	
PD	09-AUG-2001.
XX	
PF	29-JAN-2001; 2001WO-US00661.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-476286/51.
XX	
PT	Novel single exon nucleic acid probe used to measuring gene expression
PT	in a human breast -
XX	
PS	Claim 25; SEQ ID NO 3335; 322pp; English.

CC The present invention relates to novel single exon nucleic acid probes.
CC
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
Sequence 378 BP: 80 A; 128 C; 97 G; 73 T; 0 other:

Query Match	20.4%	Score 351	DB 22	Length 378
Best Local Similarity	99.7%	Pred. No. 6.9e-83		
Matches 362	Conservative 0	Mismatches 0	Indels 1	Gaps 1
Qy 1	ctgtgccccctcagatcaagaagcctcttcctctgtgtgtgcacaagtgatcggagac	60		
Db 362	CCTGGCCCCCTCGATCAAGAACGCCCTTCTTCTCTGGGGCAACGGTcCGGGCAC	303		
Qy 61	ccctctatgagacagcaagaagcagagctltgtggttgaagagagcttggaggtgaag	120		
Db 302	CCCTCTATGGGACAGCAAGAACACAGAGCTTTGTGGGTGAGAGAGCTGGAGGTGAAG	243		
Qy 121	ggaagatgtagaagagtgtagaggggagatcctgtacaggtgtgtctgtggctgtatc	180		
Db 242	GGAGATGAGAGGGAGGTGAGGGGGGATCTTGTACGGCTTCTTGGGGCTGATCTCGATAT	183		
Qy 181	accacaagcttgagctcttgaagcccaagcccaaggggacaggtgtgaagaaagttcatcc	240		
Db 182	ACCACAACTGGTGCTTCAGGGCCAAACCCAGCAGGGGGCCAGGATGTAGAGAAAGTCCATCC	123		
Qy 241	ggaagcttgcagtgagccagcttggagagcccttgggctcacaattcccccatctgtgaagcgc	300		
Db 122	GGAGTGTGTCATGTGGCCAGCTGGGAGACCCCTGGGGCTCATTTTCCCATCTGTGTGACCCCT	63		
Qy 301	atgacacagctgacacattccactccgcgtactgcatagtgcctgtgacatagtgctaaagg	360		
Db 62	ATGACACACTGACACCTTTCACCTCCGCTACTGATGATGGCCCTGTG-CATAGTGTGTAAGG	4		

OY 361 agc 363
 111
 DB 3 AGC 1

RESULT 10
 ABA08485
 ID ABA08485 standard; cDNA: 547 BP.
 AC ABA08485:
 DT 11-JAN-2002 (first entry)
 XX

Human AMP-activated protein kinase subunit homologue cDNA, SEQ ID NO:261.
 Human: cytokine; cell proliferation; cell differentiation; growth factor;
 haematopoiesis regulation; tissue growth; immunomodulator; activin;
 inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
 proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 chronic inflammatory condition; proliferative retinopathy;
 atherosclerosis; coronary heart disease; arterial ischaemia;
 bone disorder; osteoporosis; vascular growth disorder;
 tissue regeneration; wound healing; infection; immune disorder;
 cell culture; drug screening; gene therapy; antiinflammatory;
 antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KM cytostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
 KM antifungal; vulnery; antitumor; ss.

OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 XX
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457740/49.
 XX
 DR P-PSDB: ABB11241.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 1: Page 429; 1963pp; English.
 XX
 CC Sequences ABB10991-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical

CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX SQ Sequence 547 BP; 112 A; 172 C; 133 G; 130 T; 0 other;

Query Match 11.4%; Score 196; DB 22; Length 547;

Best Local Similarity 95.3%; Pred No 7,3e-42;

Matches 202; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1510 ccacccaacaggtccctcgtcgtcccgccctcctcctacacgaataccaaga 1569
 DB 272 cctgcacatccttggtccctcgtcccgccctcctcctacacgaataccaaga 331
 OY 1570 ttgggacatcgacatcccgagacttgctgtgctggagagaagacacccatccgac 1629
 DB 332 ttgggacatcgacatcccgagacttgctgtgctggagagaagacacccatccgac 391
 OY 1630 tgcactggaacatcttgtagaccggtgtgtcgtcactgacctggtgcaagaatgtgg 1689
 DB 392 tgcactggaacatcttgtagaccggtgtgtcgtcactgacctggtgcaagaatgtgg 451
 OY 1690 taccaccccccaagatgagaggtcgggctgg 1721
 DB 452 taccaccccccaagatgagaggtcgggctgg 483

RESULT 11
 AAH43685
 ID AAH43685 standard; cDNA: 1647 BP.
 AC AAH43685:
 DT 21-JAN-2002 (first entry)
 XX
 DE PRKAG3 cDNA.
 XX
 KW Human: AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
 KW metabolic disease; diabetes; obesity; substitution; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 20..1489
 FT /*tag= a
 FT /product= "PRKAG3"
 FT 230
 FT /*tag= b
 FT /label= "C230G"
 FT /note= "Causes P71A"
 FT 559
 FT /*tag= c
 FT /label= "T559C"
 FT /note= "Silent variation"
 FT 1037
 FT /*tag= d
 FT /label= "C1037T"

```

FT      /note- "Causes R340W"
XX      MO200177305-A2.
PN      18-OCT-2001.
XX      06-APR-2001: 2001WO-SF00765.
XX      PF      07-APR-2000: 2000US-195655P.
XX      PA      (AREX-) AREXIS AB.
XX      PI      Andersson L, Luthman H, Marklund S;
DR      WPI: 2001-657170/75.
XX      P-PDB: Q0B47679.
XX      PT      New variants of human AMP-activated protein kinase gamma3 subunit
PT      associated with a metabolic disease e.g. diabetes or obesity and method
PT      for determining a risk estimate of diseases in subject by detecting the
PT      variant -
PS      Disclosure: Fig 5: 25pp: English.
XX      CC      This sequence represents the full length cDNA encoding the human
CC      AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC      the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC      a risk estimate of a metabolic disease, such as diabetes or obesity,
CC      in a subject. The variation may occur in exons 3, 4 or 10. In exon
CC      3 a variation may be a substitution of a G for a C at nucleotide 320,
CC      resulting in the amino acid substitution P71A; in exon 4 a variation may
CC      be a substitution of a T for a C at nucleotide 550; and in exon 10
CC      a variation may be a substitution of a T for a C at nucleotide 1037,
CC      resulting in the amino acid substitution R340W. There may also be
CC      nucleotide variation in intron 6. The numbering of these
CC      variations is based on the full length cDNA as given, rather than on
CC      position 1 of the open reading frame.
XX      SQ      Sequence 1647 BP: 346 A; 502 C; 462 G; 337 T; 0 other;

Query Match          9.8%: Score 168.2; DB 22: Length 1647;
Best Local Similarity 95.6%; Pred. No. 2.5e-34;
Matches 173; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1510 ccactcaaccaggctccctcgtcgcccgcgccccctcctcaacgaatccaaga 1569
DB      1009 cctgcacaccttggcttcctcgtcgcccgcgccccctcctcaacgaatccaaga 1068
OY      1570 ttggagcatcgagcacattccgagacttgctgtgtctggaagacaagccatcctgac 1629
DB      1069 ttggagcatcgagcacattccgagacttgctgtgtctggaagacaagccatcctgac 1128
OY      1630 tgcactgacacatcttgtgagccgcggtgctgtcactgctgtgtcaacgaatgtgg 1689
DB      1129 tgcactgacacatcttgtgagccgcggtgctgtcactgctgtgtcaacgaatgtgg 1188
OY      1690 t 1690
DB      1189 t 1189

RESULT 12
AAD03296
ID      AAD03296 standard; DNA: 2109 BP.
XX      AC      AAD03296:
XX      DT      13-JUN-2001 (first entry)
XX      DE      Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX      HU      Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
```

KM	PRKAG3, diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW	genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KM	cystathione beta synthase; CBS; cardiact; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	5'UTR
FT	1..471
FT	/tag= a
FT	CDS
FT	472..1389
FT	/tag= b
FT	/product= "Human Prkag3 protein"
FT	1390..2109
FT	/tag= c
XX	
PX	WO200120003-A2.
PD	
PD	22-MAR-2001.
XX	
PF	11-SEP-2000; 2000WO-EP09896.
PX	
PX	10-SEP-1999; 99EP-0402236.
PR	18-MAY-2000; 2000EP-0401388.
XX	
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.
PA	(ANDE/) ANDERSSON L.
PA	(LOOF/) LOOFT C.
PA	(KALM/) KALM E.
XA	
PI	Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI	Iannucelli N, Gellin J, Le Roy P, Chardon P;
XX	
DR	WPI: 2001-244810/25.
XX	P-PSDB: AAEO0221.
PT	New variants of the gamma subunit of vertebrate adenosine
PT	monophosphate-activated kinase for diagnosis or treatment of disorders
PT	associated with energy metabolism such as diabetes, obesity, and
PT	myopathy -
XX	
PS	
XX	
XX	Claim 12; Fig 2; 71pp; English.
CC	The present sequence is a cDNA encoding human adenosine monophosphate
CC	(AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC	(PRAK3). Mutation in Prkag3 results in an altered regulation of
CC	carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC	useful as therapeutic for treating carbohydrate metabolism disorders such
CC	as diabetes, obesity, and disorders associated with muscle metabolism
CC	such as myopathy and cardiovascular diseases, to modulate AMPK
CC	activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC	and its functionally altered mutants are useful for the diagnostic
CC	evaluation, genetic testing and prognosis of a metabolic disorder,
CC	preferably a carbohydrate metabolism disorder. Primers that can detect
CC	a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC	useful for detecting a dysfunction of carbohydrate metabolism resulting
CC	from the expression of a functionally altered allele of PRKAG3.
CC	Transgenic animal and host cell transformed with PRKAG3 or a
CC	heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC	screening compounds able to modulate AMPK activity. Nucleic acid
CC	encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC	in a sequence encoding the first cystathione beta synthase (CBS) domain
CC	of PRKAG3 and is useful in gene therapy.
XX	
SO	Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;

Query Match 9.8%; Score 168.2; DB 22; Length 2109;
 Best Local Similarity 95.6%; Pred. No. 2.8e-34;
 Matches 173; Conservative 0; Mismatches 8; Indels 0; Gaps 0

```

0y   1510 ccattctaacagaagtccctcgtcgccccgcaccttctctacgacatccaaga 1569
      |||
Db   909 ccggcaatcttggttcctctgctgccgcgcctctctctcacgcacatccaaga 968
  
```


|||||
Db 1035 cgcactggagacatctctcgtggaccgcgctgtgtcctcgcgtgcctcgtgtccacgaaactgg 1094

Search completed: October 3, 2002, 16:30:16
Job time: 14286 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic acid - nucleic acid search, using sw model

Run on: October 3, 2002, 11:39:20 ; Search time 180.77 Seconds
(without alignments)
2339.884 Million cell updates/sec

Title: US-09-826-581-3

Perfect score: 1722
Sequence: 1 cctggccctcagatcaaga.....gatgagagctcggctcgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTODS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	68.8	4.0	1576 3	US-09-101-146-63 Sequence 63, Appl
2	67.6	3.9	7218 1	US-08-232-463-14 Sequence 14, Appl
3	63.8	3.7	1435 2	US-08-878-989-14 Sequence 14, Appl
4	63.8	3.7	1435 4	US-09-272-796-14 Sequence 14, Appl
5	61	3.5	7218 1	US-08-232-463-14 Sequence 14, Appl
6	39.4	2.3	289 4	US-09-007-005-17 Sequence 17, Appl
7	39.4	2.3	289 4	US-09-244-796-17 Sequence 17, Appl
8	39	2.3	53526 3	US-08-658-136-2 Sequence 2, Appl
9	39	2.3	53577 3	US-08-658-136-1 Sequence 1, Appl
10	38.6	2.2	320 4	US-09-165-264-7 Sequence 7, Appl
11	38.6	2.2	320 4	US-09-165-264-14 Sequence 14, Appl
12	37.4	2.2	319 4	US-09-165-264-8 Sequence 8, Appl
13	36.2	2.1	320 4	US-09-165-264-13 Sequence 13, Appl
14	36	2.1	320 4	US-09-165-264-11 Sequence 11, Appl
15	36	2.1	2278 1	US-07-959-943-6 Sequence 6, Appl
16	36	2.1	2415 1	US-07-959-943-8 Sequence 8, Appl
17	35.8	2.1	12001 1	US-08-458-568A-11 Sequence 11, Appl
18	35.4	2.1	318 4	US-09-165-264-12 Sequence 12, Appl
19	35.2	2.0	4066 3	US-09-215-966-21 Sequence 21, Appl
20	34.4	2.0	707 4	US-08-998-416-876 Sequence 876, Appl
21	34.4	2.0	152331 3	US-09-128-155-16 Sequence 16, Appl
22	34.2	2.0	6306 1	US-08-466-390-3 Sequence 3, Appl
23	34.2	2.0	6306 1	US-08-470-950-3 Sequence 3, Appl
24	34.2	2.0	6306 1	US-08-467-781-3 Sequence 3, Appl
25	34.2	2.0	6306 1	US-08-195-487-3 Sequence 3, Appl
26	34.2	2.0	6306 2	US-08-483-924-3 Sequence 3, Appl
27	34.2	2.0	6306 5	PCT-US93-06160-3 Sequence 3, Appl

C 28	34.2	2.0	7705 2	US-08-687-080-115 Sequence 115, Appl
C 29	34	2.0	2117 4	US-09-155-768-11 Sequence 11, Appl
C 30	34	2.0	6232 4	US-08-456-200B-11 Sequence 11, Appl
C 31	34	2.0	7175 1	US-08-455-543A-8 Sequence 8, Appl
C 32	34	2.0	7175 2	US-08-193-078B-8 Sequence 8, Appl
C 33	34	2.0	7175 2	US-08-223-305C-8 Sequence 8, Appl
C 34	34	2.0	7175 2	US-08-149-097D-8 Sequence 8, Appl
C 35	34	2.0	7175 3	US-08-949-386-8 Sequence 8, Appl
C 36	34	2.0	7175 3	US-08-450-562-8 Sequence 8, Appl
C 37	34	2.0	7175 4	US-08-984-709A-8 Sequence 8, Appl
C 38	34	2.0	7177 4	US-09-268-163-7 Sequence 7, Appl
C 39	34	2.0	7266 3	US-08-713-118-1 Sequence 1, Appl
C 40	34	2.0	7266 3	US-09-452-007-1 Sequence 1, Appl
C 41	34	2.0	7362 1	US-08-455-543A-7 Sequence 7, Appl
C 42	34	2.0	7362 2	US-08-193-078B-7 Sequence 7, Appl
C 43	34	2.0	7362 2	US-08-223-305C-7 Sequence 7, Appl
C 44	34	2.0	7362 2	US-08-149-097D-7 Sequence 7, Appl
C 45	34	2.0	7362 3	US-08-949-386-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-101-146-63
Sequence 63, Application US/09101146
Patent No. 6124125
GENERAL INFORMATION:
APPLICANT: Dartmouth College, St. Vincent's Institute of
APPLICANT: Medical Research, Kemp et al.
TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/101,146
FILING DATE: October 7, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PN7450
FILING DATE: 8 JAN 1996
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: DC-0050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 1576
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-09-101-146-63

Query Match 4.0%; Score 68.8; DB 3; Length 1576;
Best Local Similarity 67.4%; Pred. No. 4.2e-09;
Matches 97; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 1240 cagctgtttgaagctgtctcaacccctcaagaacgagatcgcgtctgttct 1299

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

```

RESULT      6
US-09-007-005-17
: Sequence 17, Application US/09007005B
: Patent No. 6258558
: GENERAL INFORMATION:
: APPLICANT: Szostak, Jack W.
: APPLICANT: Roberts, Richard W.
: APPLICANT: Liu, Rihc
: TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
: TITLE OF INVENTION: FUSIONS
: FILE REFERENCE: 00786/35003
: CURRENT APPLICATION NUMBER: US/09/007,005B
: CURRENT FILING DATE: 1998-01-14
: EARLIER APPLICATION NUMBER: 60/035,963
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 60/064,491
: EARLIER FILING DATE: 1997-11-06
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 289
: TYPE: RNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Translation template
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(289)
: OTHER INFORMATION: n = A,T,C or G
: US-09-007-005-17

```

Query Match 2.3%; Score 39.4; DB 4; Length 289;
Best Local Similarity 2.7%; Pred. No. 0.16;
Matches 6; Conservative 107; Mismatches 110; Gaps 0;

[illegible]

```

::: : :: : :| :|: :: : :| : : : ::
Db 203 rnrnrnrnrnrnrsrcrgrcrurgrcgrurararcrur 245

```

```

RESULT 7
US-09-244-796-17
Sequence 17, Application US/09244796
Patent No. 6281344
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihue
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
TITLE OF INVENTION: FUSIONS
FILE REFERENCE: 00786/350007
CURRENT APPLICATION NUMBER: US/09/244,796
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/0035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
EARLIER FILING DATE: 1997-11-06
EARLIER APPLICATION NUMBER: 09/007,005
EARLIER FILING DATE: 1996-01-14
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
FEATURE: Translation template
OTHER INFORMATION:
FEATURE:
NAME/key: misc_feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

```

RESULT      8
US-08-658-136-2/c
? Sequence 2, Application US/08658136
? Patent No. 6071717
?
? GENERAL INFORMATION:
?
? APPLICANT : KLINGER, KATHERINE W
?
? APPLICANT : LANDERS, GREGORY M
?
? APPLICANT : BURR, TIMOTHY C
?
? APPLICANT : CONNORS, TIMOTHY D
?
? APPLICANT : DACKOWSKI, WILLIAM
?
? APPLICANT : GERMINO, GREGORY
?
? APPLICANT : QIAN, FENG
?
? TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
?
? NUMBER OF SEQUENCES: 58

```


[illegible]

RESULT 11
US-09-165-264-14
; Sequence 14, Application US/09165264
; Patent No. 6107510

```

? APPLICANT: Vinayagamorthy, Thirumalaiah
? TITLE OF INVENTION: Multi-LoCI Genomic Analysis
? FILE REFERENCE: 44747
? CURRENT APPLICATION NUMBER: US/09/165,264
? CURRENT FILING DATE: 1998-10-01
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 14
? LENGTH: 320
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
? US-09-165-264-14

```

Query Match	2.28	Score 38.6	DB 4	Length 320
Best Local Similarity	48.88	Pred. No. 0.27		
Matches 104	Conservative 0	Mismatches 109	Indels 0	Gaps 0

[illegible]

RESULT 12
US-09-165-264-8
: Sequence 8, Application US/09165264
: Patent No. 6197510

```

1  APPLICANT: Vinayagamorthy, Thirulayah
2  TITLE OF INVENTION: Multi-loci Genomic Analysis
3  FILE REFERENCE: 44747
4  CURRENT APPLICATION NUMBER: US/09/165,264
5  CURRENT FILING DATE: 1998-10-01
6  NUMBER OF SEQ ID NOS: 14
7  SOFTWARE: PatentIn Ver. 2.1
8  SEQ ID NO 8
9  LENGTH: 319
10 TYPE: DNA
11 ORGANISM: Artificial Sequence
12 FEATURE:
13 OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
14 US-09-165-264-8

```

Query Match	2.28;	Score 37.4;	DB 4;	length 319;
Best Local Similarity	50.38;	Pred. No. 0.56;		
Matches	92;	Conservative	0;	Mismatches 91;
			Indels	0;
			Gaps	0;

Qy	777	ggagagagcgcgacgcggaggagagaaaccacgaagaagcgccgtctcgttcgatctcggcacgaagc	836
Db	130	ggag	189
Qy	837	taaagctcggagagatcggcacagtctggagagatgcctcgcagagagagagagacaataaagagac	896
Db	190	ggag	249
Qy	897	cctcggctgcctcgcagcggaaagagaaagctgcctcggagacatgcagaagctgaagcaagctgaacgctt	956
Db	250	ggag	309
Qy	957	gcc gcc	
Db	310	ccc c12	

RESULT 13
US-09-165-264-13
; Sequence 13, Application US/09165264
; Patent No. 6107510

```

? APPLICANT: Vinayagamoorthy, Thuralayah
? TITLE OF INVENTION: Multi-Loci Genomic Analysis
? FILE REFERENCE: 44747
? CURRENT APPLICATION NUMBER: US/09/165,264
? CURRENT FILING DATE: 1998-10-01
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 13
? LENGTH: 320
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
? US-09-165-264-13

```

Query Match	2.1%	Score 36.2;	DB 4;	Length 320;
Best Local Similarity	49.28;	Pred. NO. 1.2;		
Matches	95;	Conservative	0;	Mismatches 98;
			Indels	0;
			Gaps	0

[illegible]

RESULT 14
US-09-165-264-11
; Sequence 11, Application US/09165264
; Patent No. 6107510

```
: APPLICANT: Vinayagamorthy, Thuralayah  
: TITLE OF INVENTION: Multi-Loci Genomic Analysis  
: FILE REFERENCE: 44747  
: CURRENT APPLICATION NUMBER: US/09/165,264  
: CURRENT FILING DATE: 1998-10-01
```

```
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11
```

```
Query Match      2.1% Score 36; DB 4; Length 320;
Best Local Similarity 50.6%; Pred. No. 1.3;
Matches 87; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 777 ggaagagctgagctggagaggaacccgaaagggctgtgtgtgagccagggct 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 ggggggggggggggggggggggggggggggggggggggggggggggggggggg 134

QY 837 taagggtgagagtgagcagtgaggatgctctgagtgaggaagggggaacaataggagc 896
    || || || || || || || || || || || || || || || || || || ||
Db 195 ggggggggggggggggggggggggggggggggggggggggggggggggggggg 254

QY 897 ctccgggtccctcagcgaaggaagctgcctcggagctgcaagtgagcaggt 948
    ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 255 ggggggggggggggggggggggggggggggggggggggggggggggggggggg 306
```

RESULT 15

US-07-959-943-6/c

```
; Sequence 6, Application US/07959943
; Patent No. 5418162
; GENBANK INFORMATION:
; APPLICANT: Blakely, Randy D.
; APPLICANT: Fremieu Jr., Robert T.
; TITLE OF INVENTION: Serotonin Transporter CDNA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park
; ADDRESSEE: and
; STREET: Post Office Drawer 31107
; CITY: Raleigh
; STATE: No. 5418162th Carolina
; COUNTRY: U.S.A.
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,943
; FILING DATE: 19921014
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405.38a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; POSITION IN GENOME:
; UNITS: 2278 basepairs
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 48..1868
US-07-959-943-6
```

```
Query Match      2.1% Score 36; DB 1; Length 2278;
Best Local Similarity 51.9%; Pred. No. 3;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 611 ggtgagagctgggcttggaaatcttatgagcaccagagggcggggagggagtc 670
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1943 GGGGAGGAGTCAGCTGGGCTTTTCAGAGCTGAGAGAGCTGCGGAAGAGCTGCTC 1884

QY 671 tccaggagcctgtgtccctcagaagcccaagcttcttgacttcaggagtcgtagt 730
    || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1883 TCCAGAGGTCTGTACACAGCATTCATCCGATGTCGCCACACAGGATTTCTGTGGCT 1824

QY 731 ctctaagtcagaatcctatgagaltgaacaacataag 766
    || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1823 TTCAGGAGTGATCTTTTATATATGCGCTCTTAAG 1788
```

Search completed: October 3, 2002, 16:18:58
Job time: 16778 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:46:08 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700

Perfect score: 101

Sequence: 1 aggtccccctgtgtgaggaag.....tggtgccctagaagccacag 101

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb.ba:*
2: gb.hlg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pal:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.hlg.hum:*
31: em.hlg.inv:*
32: em.hlg.other:*
33: em.hgo.inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	101	100.0	1722	6	AX281580	Sequence	AX281580
c	2	101	100.0	196554	2	AC073128	AC073128 Homo sapi
c	3	101	100.0	206854	9	AC009974	AC009974 Homo sapi
c	4	99.4	98.4	152129	2	AC027416	AC027416 Homo sapi
c	5	35	34.7	2539	4	U01206	U01206 Bos taurus
c	6	35	34.7	127794	4	AL138688	AL138688 Human DNA
c	7	33.2	32.9	692	5	AF461698	AF461698 Gallus ga
c	8	33.2	32.9	895	5	S69088	S69088 gHox-homoe
c	9	33.2	32.9	1084	5	CHKPRX1	D13433 Gallus gall
c	10	32.6	32.3	6335	9	D86982	D86982 Human mRNA
c	11	32.6	32.3	54018	2	AC091776	AC091776 Chlamydom
c	12	32.6	32.3	65898	2	AC090435	AC090435 Chlamydom
c	13	32.6	32.3	161999	2	AL133352	AL133352 Homo sapi
c	14	32.6	32.3	172748	2	AC097417	AC097417 Rattus no
c	15	32.6	32.3	203200	9	AC008744	AC008744 Homo sapi
c	16	32.4	32.1	3761	9	HSMB00653	AL080141 Homo sapi
c	17	32.4	32.1	4648	9	AF274863	AF274863 Homo sapi
c	18	32.4	32.1	5888	9	AF214521	AF214521 Sus scrof
c	19	32	31.7	45375	5	AF328738	AF328738 Agelaius
c	20	32	31.7	73699	2	AC094322	AC094322 Rattus no
c	21	31.8	31.5	104514	2	AC008713	AC008713 Homo sapi
c	22	31.8	31.5	135354	2	AC106731	AC106731 Homo sapi
c	23	31.8	31.5	136657	2	AL669924	AL669924 Homo sapi
c	24	31.8	31.5	171428	2	AC032044	AC032044 Homo sapi
c	25	31.6	31.3	171428	9	AC013407	AC013407 Homo sapi
c	26	31.6	31.3	175999	2	AC021154	AC021154 Homo sapi
c	27	31.6	31.3	183491	9	CNS010U4	AL133163 Human chr
c	28	31.4	31.1	63937	2	AC024290	AC024290 Homo sapi
c	29	31.4	31.1	67600	2	AC105210	AC105210 Homo sapi
c	30	31.4	31.1	88928	2	AC106624	AC106624 Rattus no
c	31	31.4	31.1	157790	2	AL357509	AL357509 Homo sapi
c	32	31.4	31.1	164959	2	AC087823	AC087823 Homo sapi
c	33	31.4	31.1	207636	2	AC019122	AC019122 Homo sapi
c	34	31.4	31.1	215692	2	AC090977	AC090977 Mus muscu
c	35	31.4	31.1	215727	2	AL358795	AL358795 Homo sapi
c	36	31.2	30.9	145556	2	OSJN00092	AL606651 Oryza sat
c	37	31.2	30.9	154140	2	AC010758	AC010758 Homo sapi
c	38	31.2	30.9	157437	2	AC021669	AC021669 Homo sapi
c	39	31.2	30.9	167882	2	AC096852	AC096852 Sus scrof
c	40	31.2	30.9	174281	2	AC068338	AC068338 Homo sapi
c	41	31.2	30.9	177433	9	AC011747	AC011747 Homo sapi
c	42	31.2	30.9	186229	2	AC095022	AC095022 Sus scrof
c	43	31.2	30.9	210651	2	AC105137	AC105137 Homo sapi
c	44	31	30.7	1105	6	AX090307	AX090307 Sequence
c	45	31	30.7	1105	9	AF305687	AF305687 Homo sapi

ALIGNMENTS

RESULT	1	AX281580	1722 bp	DNA	linear	PAT 03-NOV-2001
LOCUS	AX281580	Sequence 3 from Patent WO0177305.				
DEFINITION	AX281580					
ACCESSION	AX281580					
VERSION	AX281580.1	GI:16608831				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.				
AUTHORS		Andersson, L., Luthman, H. and Marklund, S.				
TITLE		Varicants of the human amp-activated protein kinase gamma 3 subunit				
JOURNAL		Patent: WO 0177305-A 3 18-OCN-2001; Arexis AB (SE)				

FEATURES
source
1..1722
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT
ORIGIN
321 a 504 c 534 g 363 t

Query Match 100.0% Score 101: DB 6: Length 1722:
Best Local Similarity 100.0%: Pred. No. 1.3e-16:
Matches 101: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Qy 1 aggtccccctgttaagagctgagctgggaatctatagcaccagagagggagggcg 60
|||||
Db 600 AGGTCGCCCCCTGTCAGAGTCGCGCTGGCAATCTTATGGCACCCAGAGGGCGGGCGC 659
|||||

Qy 61 gaaggagagtcctcctgagagctgtgtgccatgaagcccaag 101
|||||
Db 660 GAGCGGAGTCCCTCGAGAGCTGTCGCCCTAGAACCCACG 700
|||||

RESULT 2
AC073128/c
LOCUS AC073128 196554 bp DNA linear HTG 21-FEB-2001
DEFINITION Homo sapiens chromosome 2 clone RP11-64705, WORKING DRAFT SEQUENCE,
17 unordered pieces.
AC073128
AC073128.3 GI:13027579
KEYWORDS HTG: HTGS_PHASE1: HTGS_DRAFT: HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 196554)
WATERSTON, R. H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 196554)
WATERSTON, R. H.
Direct Submission
Submitted (08-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 21, 2001 this sequence version replaced gi:8469048.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0647005
----- Summary Statistics -----
Sequencing vector: M13: 98%
Sequencing vector: plasmid: 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187795 bases at least Q40
Consensus quality: 190513 bases at least Q30
Insert size: 200000; agarose-fp
Insert size: 194954; sum-of-contigs
Quality coverage: 5.58 in Q20 bases; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1157: contig of 1157 bp in length
* 1158 1257: gap of unknown length
* 1258 3600: contig of 2343 bp in length
* 3601 3700: gap of unknown length
* 3701 5103: contig of 1403 bp in length
* 5104 5203: gap of unknown length
* 5204 8524: contig of 3321 bp in length

8525 8624: gap of unknown length
* 8625 11856: contig of 3232 bp in length
* 11857 11956: gap of unknown length
* 11957 15783: contig of 3827 bp in length
* 15784 15883: gap of unknown length
* 15884 21906: contig of 6023 bp in length
* 21907 22005: gap of unknown length
* 22007 28887: contig of 6881 bp in length
* 28888 35255: gap of unknown length
* 35256 35355: gap of 6268 bp in length
* 35356 44642: contig of 9287 bp in length
* 44643 44743: gap of unknown length
* 44743 58275: contig of 13533 bp in length
* 58276 58375: gap of unknown length
* 58376 73816: contig of 15441 bp in length
* 73817 73916: gap of unknown length
* 73917 92140: contig of 18224 bp in length
* 92141 92240: gap of unknown length
* 92241 113337: contig of 21097 bp in length
* 113338 113437: gap of unknown length
* 113438 130325: contig of 16888 bp in length
* 130326 130425: gap of unknown length
* 130426 149287: contig of 18862 bp in length
* 149288 149388: gap of unknown length
* 149389 196554: contig of 47167 bp in length.

FEATURES
source
1. 196554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-64705"
1. 1157
/note="assembly_name:Contig17"
1258. 3600
/note="assembly_name:Contig18"
3701. 5103
/note="assembly_name:Contig19
clone_end:77
vector_side:right"
5204. 8524
/note="assembly_name:Contig20"
8625. 11856
/note="assembly_name:Contig21"
11957. 15783
/note="assembly_name:Contig22"
15884. 21906
/note="assembly_name:Contig23"
22007. 28887
/note="assembly_name:Contig24"
28988. 35255
/note="assembly_name:Contig25"
35356. 44642
/note="assembly_name:Contig26"
44743. 58275
/note="assembly_name:Contig27"
58376. 73816
/note="assembly_name:Contig28"
73917. 92140
/note="assembly_name:Contig29"
92241. 113337
/note="assembly_name:Contig30"
113438. 130325
/note="assembly_name:Contig31"
130426. 149287
/note="assembly_name:Contig32"
149388. 196554
/note="assembly_name:Contig33
clone_end:SP6
vector_side:right"

BASE COUNT 52296 a 46993 c 45889 g 49770 t 1606 others
ORIGIN

Query Match	100.0%	Score 101	DB 2	Length 196554
Best Local Similarity	100.0%	Pred. No. 4,4e-17		
Matches 101	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1 aggtcccccctggtgaagagatgagccttggaatcctatggtgcacccagagagggcgagcg 60				
Db 60519 AGCTCCCCCCCTGTCGAGAGAGTGCGCTGGGAACTTATGCGCACCCAGAGGGCGGGCG 60460				
QY 61 gaagggaaatcctcctcgtgagcctggtgacctagaaccacag 101				
Db 60459 GAGGGAGTCTCTCTCGAGACCTGCTGCCCTACAAACCCACG 60419				

RESULT	_3
LOCUS	AC009974/c
DEFINITION	AC009974
ACCSSION	Homo sapiens BAC clone RP11-459119 from 2, complete sequence.
VERSION	AC009974
KEYWORDS	AC009974..9 GI:16799058
SOURCE	HTG.
ORGANISM	human. Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 206854)
JOURNAL	Sulston,J.E. and Waterston,R. Toward a complete human genome sequence Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE	99063792
REFERENCE	2 (bases 1 to 206854)
AUTHORS	Harris,A. and Cotton,M.
TITLE	The sequence of Homo sapiens BAC clone RPL1-459I19
JOURNAL	Unpublished (2001)
REFERENCE	3 (bases 1 to 206854)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (08-SEP-1999) Genome Sequencing Center, Washington

REFERENCE
AUTHORS 4 (bases 1 to 206854)
TITLE Waterston, R. H.
JOURNAL Direct Submission
Submitted (08-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 206854)
AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

6 (bases 1 to 206854)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 8, 2001 this sequence version replaced qi:13411203.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0459119

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., paired quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCR1 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Fritgen, E., Tatenoe, M., Catenease, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.rsgen.com>) or Pletier de Jong and co-workers at the Roswell Park Cancer Institute (<http://dabcpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-1077K22; the clone sequenced to the right is RP11-64705. Actual start of this clone is at base position 1 of RP11-459119; actual end is at base position 206854 of RP11-459119.

Data from AC079810 and AC073128 was used to finish this clone, AC0009974. Polymorphisms have been identified between AC073128 and AC0009974. A single plasmid region exists between 38812-38903. An unresolved tandem in the HEV SVA exists between 184390-185163. PCR suggests that approximately 1700 dps are missing.

FEATURES

source	1. .206854 /organism="Homo sapiens" /db_xref="taxon: 9606" /chromosome="2" /map="2" /clone="RP11-459119" /clone_lib="RPC1-11"
misc_feature	1. .37 /note="match to EST BE314060 (NID:g1134719)"
misc_feature	1. .37 /note="match to EST BI059713 (NID:g14467240)"
misc_feature	1. .37 /note="match to EST BF183086 (NID:g11061273)"
misc_feature	1. .37 /note="match to EST AL567345 (NID:g12920610)"
misc_feature	1. .37 /note="match to EST AW880850 (NID:g8042860)"
misc_feature	1. .37 /note="match to EST BF304755 (NID:g11251653)"
misc_feature	1. .37 /note="similar to Homo sapiens EST B114348 (NID:g14565249)"
misc_feature	1. .37 /note="match to EST BG477625 (NID:g13409904)"
misc_feature	1. .37 /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature	3. .37 /note="match to EST BE908408 (NID:g10402954)"
misc_feature	4. .37 /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
misc_feature	164. .662 /note="match to EST C05773 (NID:g1502549)"
misc_feature	220. .221 /note="match to EST BG470047 (NID:g13402322)"
misc_feature	281. .344 /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature	281. .344 /note="match to EST BE908408 (NID:g10402954)"

```

misc_feature      281..344      /note="match to EST BF304755 (NID:g11251653)"
misc_feature      281..344      /note="match to EST BF183086 (NID:g11061273)"
misc_feature      281..344      /note="similar to Homo sapiens EST B1114348
                               (NID:g14565249)"
misc_feature      281..344      /note="match to EST BG477625 (NID:g13409904)"
misc_feature      281..344      /note="match to EST B1059713 (NID:g14467240)"
misc_feature      281..344      /note="match to EST AL567345 (NID:g12920610)"
misc_feature      281..344      /note="match to EST BE314060 (NID:g9134719)"
misc_feature      281..344      /note="match to EST A1670836 (NID:g4850567) wa04g10.x1"
misc_feature      281..344      /note="match to EST AW880850 (NID:g8042860)"
misc_feature      294..344      /note="match to EST BG470047 (NID:g13402322)"
misc_feature      588..929      /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature      594..763      /note="match to EST T64073 (NID:g667938) yc05d12.r1"
misc_feature      594..763      /note="match to EST BF304755 (NID:g11251653)"
misc_feature      594..763      /note="match to EST BG470047 (NID:g13402322)"
misc_feature      594..763      /note="match to EST AW880850 (NID:g8042860)"
misc_feature      594..731      /note="match to EST BE314060 (NID:g9134719)"
misc_feature      594..764      /note="match to EST B1059713 (NID:g14467240)"
misc_feature      594..763      /note="match to EST BG477625 (NID:g13409904)"
misc_feature      594..742      /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature      594..763      /note="match to EST BE908408 (NID:g10402954)"
misc_feature      594..763      /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature      594..763      /note="match to EST A1670836 (NID:g4850567) wa04g10.x1"
misc_feature      594..763      /note="match to EST BF183086 (NID:g11061273)"
misc_feature      594..763      /note="similar to Homo sapiens EST B1114348
                               (NID:g14565249)"
misc_feature      599..763      /note="match to EST AL567345 (NID:g12920610)"
misc_feature      617..1084      /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
misc_feature      622..763      /note="match to EST A1860958 (NID:g5514574) w156f05.x1"
misc_feature      622..763      /note="similar to Mus musculus EST A1196847 (NID:g3749453)
                               u167e01.x1"
misc_feature      684..763      /note="match to EST BG992568 (NID:g14396638)"
misc_feature      962..1084      /note="match to EST A1656812 (NID:g4740791) tt54b06.x1"
misc_feature      967..1084      /note="match to EST BE908408 (NID:g10402954)"
misc_feature      967..1085      /note="match to EST BF304755 (NID:g11251653)"
misc_feature      967..1091      /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature      967..1071
Query Match      100.0%; Score 101; DB 9; Length 206854;
Best Local Similarity 100.0%; Pred. No. 4,3e-17;

```

```

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 aggtcccccgtgtgaggaagtgagctgggaatctataggacccagaggcgggcgcg 60
|||||
Db 166099 AGTCCCTCCCTGTCAGAGTGGCGCGGAATCTATGCGCACCCAGAGCGCGCGCGC 166040
Qy      61 gaggggagtcctccctggagcctggtgcttaagaagccacg 101
|||||
Db 166039 GAGGGGAGTCTCTCGAGACCTGTCCTCTAGAACCCACG 165999

RESULT 4
AC027416/c 152129 bp DNA linear HTG 07-JUN-2000
LOCUS Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
DEFINITION
unordered pieces.
ACCESSION AC027416
VERSION AC027416.2 GI:8317289
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 152129)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens, clone RP11-504G11
TITLE Unpublished
AUTHORS 2 (bases 1 to 152129)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barne,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McKern,A., McKernan,K., McPheters,R.,
Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2000 this sequence version replaced g1:7342115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7458
Center clone name: 504.G11
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135376 bases at least Q40
Consensus quality: 143264 bases at least Q30
Consensus quality: 146503 bases at least Q20
Insert size: 161000; agarose-1p
Insert size: 149029; sum-of-ctrls

```

Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1005: contig of 1005 bp in length
* 1006 1105: gap of 100 bp
* 2402: contig of 1297 bp in length
* 2403 2502: gap of 100 bp
* 2503 3823: contig of 1321 bp in length
* 3824 3923: gap of 100 bp
* 3924 5020: contig of 1097 bp in length
* 5021 5120: gap of 100 bp
* 5121 6161: contig of 1041 bp in length
* 6162 6261: gap of 100 bp
* 6262 7547: contig of 1286 bp in length
* 7548 7647: gap of 100 bp
* 7648 9983: contig of 2336 bp in length
* 9984 10083: gap of 100 bp
* 10084 12556: contig of 2473 bp in length
* 12557 12656: gap of 100 bp
* 12657 15043: contig of 2387 bp in length
* 15044 15143: gap of 100 bp
* 15144 17123: contig of 1980 bp in length
* 17124 17223: gap of 100 bp
* 17224 19466: contig of 2243 bp in length
* 19467 19566: gap of 100 bp
* 19567 21928: contig of 2362 bp in length
* 21929 22028: gap of 100 bp
* 22029 24319: contig of 2291 bp in length
* 24320 24419: gap of 100 bp
* 24420 27059: contig of 2640 bp in length
* 27060 27159: gap of 100 bp
* 27160 30170: contig of 3011 bp in length
* 30171 30270: gap of 100 bp
* 30271 33968: contig of 3698 bp in length
* 33969 34068: gap of 100 bp
* 34069 38179: contig of 4111 bp in length
* 38180 38279: gap of 100 bp
* 38280 42366: contig of 4087 bp in length
* 42367 42466: gap of 100 bp
* 42467 46365: contig of 3899 bp in length
* 46366 46465: gap of 100 bp
* 46466 51285: contig of 4820 bp in length
* 51286 51385: gap of 100 bp
* 51386 55871: contig of 4486 bp in length
* 55872 55971: gap of 100 bp
* 55972 60595: contig of 4624 bp in length
* 60596 60695: gap of 100 bp
* 60696 66595: contig of 5900 bp in length
* 66596 66695: gap of 100 bp
* 66696 73218: contig of 6523 bp in length
* 73219 73318: gap of 100 bp
* 73319 77115: contig of 3797 bp in length
* 77116 77215: gap of 100 bp
* 77216 85022: contig of 7807 bp in length
* 85023 85122: gap of 100 bp
* 85123 93314: contig of 8192 bp in length
* 93315 93414: gap of 100 bp
* 93415 101193: contig of 7779 bp in length
* 101194 101293: gap of 100 bp
* 101294 113090: contig of 11797 bp in length
* 113091 113190: gap of 100 bp
* 113191 123496: contig of 10306 bp in length
* 123497 123596: gap of 100 bp
* 123597 137837: contig of 14241 bp in length
* 137838 137937: gap of 100 bp

FEATURES * 137938 152129: contig of 14192 bp in length.

source
Location/Qualifiers
1. .152129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-504G11"
/clone_1lib="RPC1-11 Human Male BAC"
1. .1005
/note="assembly_fragment"
1106. .2402
/note="assembly_fragment"
2503. .3823
/note="assembly_fragment"
3924. .5020
/note="assembly_fragment"
5121. .6161
/note="assembly_fragment"
6262. .7547
/note="assembly_fragment"
7648. .9983
/note="assembly_fragment"
10084. .12556
/note="assembly_fragment"
12657. .15043
/note="assembly_fragment"
15144. .17123
/note="assembly_fragment"
17224. .19466
/note="assembly_fragment
clone_end:SP6
vector_slide:left"
19567. .21928
/note="assembly_fragment"
22029. .24319
/note="assembly_fragment"
24420. .27059
/note="assembly_fragment"
27160. .30170
/note="assembly_fragment"
30271. .33968
/note="assembly_fragment"
34069. .38179
/note="assembly_fragment"
38280. .42366
/note="assembly_fragment"
42467. .46365
/note="assembly_fragment"
46466. .51285
/note="assembly_fragment"
51386. .55871
/note="assembly_fragment"
55972. .60595
/note="assembly_fragment"
60696. .66595
/note="assembly_fragment"
66696. .73218
/note="assembly_fragment"
73319. .77115
/note="assembly_fragment"
77216. .85022
/note="assembly_fragment"
85123. .93314
/note="assembly_fragment"
93415. .101193
/note="assembly_fragment"
101294. .113090
/note="assembly_fragment"

Query Match 98.4% Score 99.4: DB 2: Length 152129;
Best Local Similarity 99.0% Pred.No.1.2e-16;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 agtcccccctgtgagagtggtgcgggaatctatgycaccccgagggcg99g9cg 60

|||||
 Db 36019 AGCTCCCTCGTGCAGAGTGGCTGCGCATCTTATGCGCAGCCAGAGCGCGCGCGC 35960
 Oy 61 gaaggagatccctcctgagagctgagccttaagaagccacg 101
 Db 35959 GAGCGAGTCTCTCTGAGCCTGCTGCTCCCTAGACGCCACG 35919

RESULT 5
 LOCUS U01206 2539 bp mRNA linear MAM 14-APR-1994
 DEFINITION Bos taurus GRK5 protein kinase mRNA, complete cds.
 ACCESSION U01206
 VERSION U01206.1 GI:437105
 KEYWORDS
 SOURCE .
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 2539)
 Premont,R.T., Koch,W.J., Inglesse,J. and Lefkowitz,R.J.
 Identification, purification and characterization of GRK5, a member
 of the family of G protein-coupled receptor kinases
 JOURNAL J. Biol. Chem. 269, 6832-6841 (1994)
 MEDLINE 94165084
 REFERENCE 2 (bases 1 to 2539)
 Premont,R.T.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1993) Richard T. Premont, Department of Medicine,
 Duke University Medical Center, Durham, NC 27710, USA
 FEATURES
 source Location/Qualifiers
 1..2539
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /tissue_type="tongue epithelium"
 /dev_stage="adult"
 173..1944
 /evidence="experimental"
 /product="GRK5 protein kinase"
 175..1947
 /codon_start=1
 /product="GRK5 protein kinase"
 /protein_id="AA17561.1"
 /db_xref="GI:437106"
 /translation="MELENIIVANTVLLKAREGGGKRRKSKKKKEILKPPHINOCED
 LRTRIDRYCGLCKOPVGRLLFROFETRGLESYIOPLDSVAEYTPDEKIGENG
 KEIMTKYLTSPKSPVITOVGRDLVSOTEEKLORCKEELPACVSHVDYRGPPIH
 VLDSMYRDLRLOKMLPROPTKMPFRORYLVLGKGGREVCACVYRATGKMTACKRLK
 KKRIKKRKGESMALNEKOILKRVNSRFVNLAVETKDALCLVLTIMGDLEFHLV
 NMGNPGEERERLALFYAEITLGLDLHHEINLVYRDLKRPENLDLDDYHIRISDLGLAV
 KIPGDLIRGVAVGVNAPVLPNNORGLSPDYWGAGLCLILEMIEGOSPGRGKERY
 KREVDRLVLETEVYSHKFSSEAKSICMLLTIDAKORLGCDEGAEEVRRHFFERN
 MNFRLEAGMLDPFVDPRAVVCADVIDIEQSTVGYNLDHDDDDFSEKSGSP
 IPQSEMIETECFEKELNVFGPHGLTSPDLNLSHPPEPKGCLLORLFRKRONNSKSS
 PNSTKSTNNHHTNSNHSNSNGSS"

BASE COUNT 673 a 641 c 712 g 513 t
 ORIGIN

Query Match 34.7% Score 35: DB 4: Length 2539:
 Best Local Similarity 59.6% Pred. No. 15;
 Matches 59: Conservative 0: Mismatches 40: Indels 0: Gaps 0:
 Oy 2 gttcccccgtgagagctgagcctgagccttaagaagccacg 61
 Db 1986 GGACCTCTCTCTTACGAGCAAGAAATTCACACAGAGAGCTCCACCTCGCGGGCGGC 2045
 Oy 62 agggagatccctcctgagagcctgagccttaagaagccac 100
 Db 2046 GGGGAAGCCCTTCAGAGACCGGGAAGAGAGCGGCCAC 2084

RESULT 6
 LOCUS AL138688/c 127794 bp DNA linear PRI 22-NOV-2000
 DEFINITION Human DNA sequence from clone RP11-264J4 on chromosome 13. Contains
 ESTs, STSs, GSSs and CpG islands. Contains the 3' part of the
 ZNF198 gene for zinc finger protein 198, a novel gene, a gene for a
 novel connexin (gap junction protein), a novel cyclophilin type
 peptidyl-prolyl cis-trans isomerase pseudogene and the GJB2 gene
 for gap junction protein beta 2, 26 kD (connexin 26), complete
 sequence.
 ACCESSION AL138688
 VERSION AL138688.27 GI:9944124
 KEYWORDS HTG; connexin; Cpg island; cyclophilin; gap junction protein; GJB2;
 prolyl cis-trans isomerase; zinc finger; ZNF198.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 127794)
 Wall,M.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Aug 29, 2000 this sequence version replaced gi:9863492.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 13, constructed by the Sanger Centre Chromosome 13
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/chr13
 RP11-264J4 is from the library RP11-11.1 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see http://bacpac.med.buffalo.edu/
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-264J4 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true right end of clone RP11-264J4 is at 127794 in this
 sequence. The true left end of clone RP11-50K3 is at 100168 in
 this sequence. The true right end of clone RP11-110K18 is at 21477
 in this sequence.

FEATURES
 source Location/Qualifiers

1..127794
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-264J4"
 /clone_1b="RP11-11.1"
 1..69
 /note="Alu/FLAM repeat: matches 2..69 of consensus"
 1..70
 /note="Alu/FLAM repeat: matches 1..69 of consensus"
 78..371
 /note="Alu repeat: matches 1..294 of consensus"
 435..508
 /note="L1MB3 repeat: matches 6109..6181 of consensus"
 941..978
 repeat_region
 repeat_region
 repeat_region
 repeat_region

```

repeat_region /note="19 copies 2 mer tt 78% conserved"
1208..1524
/note="AluSx repeat: matches 6..312 of consensus"
repeat_region 2360..2623
/note="AluSx repeat: matches 43..311 of consensus"
2764..2916
/note="MIR repeat: matches 72..260 of consensus"
3251..3534
/note="match: GSS: Em:AQ053217"
3374..3652
/note="AluYb8 repeat: matches 1..309 of consensus"
4202..4434
/note="MER33 repeat: matches 44..324 of consensus"
MER33 repeat: matches 44..324 of consensus"
4573..4877
/note="AluSc repeat: matches 1..306 of consensus"
5887..6025
/note="L1MEC repeat: matches 2194..2318 of consensus"
6026..6319
/note="Aluub repeat: matches 5..298 of consensus"
6320..6355
/note="L1MEC repeat: matches 2157..2194 of consensus"
6417..6507
/note="MER58 repeat: matches 176..269 of consensus"
6505..6556
/note="MER58 repeat: matches 5..57 of consensus"
MER58 repeat: matches 5..57 of consensus"
6568..6892
/note="L1MEC repeat: matches 2140..2148 of consensus"
join(<7550..7664,8315..8566,9190..9310,11356..12268)
/gene="ZNF198"
/note="match: CDNA5: Em:AF060181 Em:AJ224201 Em:AF012126
Em:Y13472 Em:AB002383 Em:AF161582 Em:X5808
match: ESTs: Em:AW464846 Em:AA204514 Em:BE573918
Em:AA421981 Em:BE169588 Em:AL042736 Em:AV683458
Em:BE813524 Em:AA210440 Em:AW915439 Em:BE813463
Em:BE140140 Em:AI862769 Em:BE813467 Em:AI936478
Em:AW993981 Em:BE870707 Em:AW802290 Em:AW948416
Em:AV600551 Em:BE813473 Em:AA453383 Em:AW939338
Em:BE813646 Em:BE813649 Em:AW939972 Em:BE813501 Em:W36451
Em:BE813474 Em:AW802287 Em:AW802289 Em:BE832307
Em:BE813517 Em:AW93966 Em:AA161918 Em:AW95153
Em:AW579387 Em:AA361348 Em:AI462069 Em:DB1220 Em:AI63814
Em:BE633738 Em:AA798296 Em:H72750 Em:AW573322 Em:AI112706
Em:AI642651 Em:AA851515 Em:AW565500 Em:DS9995 Em:AA926013
Em:AW535895 Em:BE140127 Em:D60911 Em:BB530161 Em:AW107660
Em:AW822686 Em:BE140127 Em:D60911 Em:BB530161 Em:AW107660
Em:BE691037 Em:AW558100 Em:AA960263 Em:AA570562 Em:C14750
Em:AV613227 Em:AI481925 Em:AV260270 Em:AI613884
Em:BE710363 Em:N71855 Em:AI606599 Em:AA251034 Em:AW193616
Em:H54287 Em:AA283078 Em:AA115047 Em:BE937769 Em:AA663837
Em:BE781058 Em:AA267676 Em:AV133871 Em:BE701954
Em:BE139491 Em:AA135989 Em:AI047555 Em:AA872477
Em:AA993582 Em:AA767826 Em:BE936168 Em:AA488120
Em:BE936169 Em:AA452155 Em:AA115537 Em:AI150524
Em:AI666742 Em:AI367945 Em:AA282776 Em:AI1381192
Em:AA860937 Em:AW243913 Em:D80504 Em:AI472479 Em:AI690992
Em:AI4354 Em:AI051311 Em:AA034499 Em:AW188335 Em:AA251580
Em:AA251581 Em:BE175951 Em:AA370561 Em:CA4338 Em:AI279353
Em:AA620081 Em:C14749 Em:BE701998 Em:AW299813 Em:DI11944
Em:AA863063 Em:BE701995 Em:AA505638 Em:AA746706 Em:N79268
Em:BE176093 Em:AA872945 Em:C14507 Em:BE934706 Em:AW301413
Em:BE701916 Em:AW885999 Em:AW902323 Em:AV613226
Em:AI866455 Em:AW068776 Em:AW901834 Em:AA253301"
/product="DA264J4.1 (zinc finger protein 198)"
/evidence="not_experimental"
7550..12268
/gene="ZNF198"
join(<7550..7664,8315..8566,9190..9310,11356..11548)
/gene="ZNF198"
/note="continues in BA110K18 (AL137119)
match: proteins: Tr:O43308 Tr:O9UBW7 Tr:O43434 Tr:O60898
Tr:O9VB55 Sw:O14202"

```

```

/codon_start=1
/evidence="not_experimental"
/product="ba264J4.1 (zinc finger protein 198)"
/protein_id="CAC16956.1"
/db_xref="GI:11322768"
/translation="YLCSRRKDNIFIDPCYOTFEDELKILRSNOPSLPDCSIFSR
VEDDYLMRIKOLGSHSPVALNTLTFPNTKRYGLKTVEDHILSLSGTVFRHMKMPLT
MEKKAKLRQVSSLCOTDNEDKITTGRKHDEDEYFEQIENAMPSCRVPVAFECYL
SKSPQNLQRMQVFLQPCSSSTDSPPVYVTSIDRLNLEMLVRLVLLKDIYDKDN
YELDEPDT"
8586..8895
/note="AluSp repeat: matches 1..308 of consensus"
9552..9843
/note="Aluio repeat: matches 4..292 of consensus"
9885..10074
/note="MER58A repeat: matches 27..224 of consensus"
10438..10804
/note="MER46B repeat: matches 1..236 of consensus"
MER46B repeat: matches 1..236 of consensus"
complement(11823..12224)
/note="match: STS: Em:G38645"
12203..12210
/gene="ZNF198"
12224
/gene="ZNF198"
12225
/gene="ZNF198"
12226
/gene="ZNF198"
12257..12262
/gene="ZNF198"
12268
/gene="ZNF198"
12269..12338
/note="5 copies 14 mer 84% conserved"
12662..12937
/note="match: GSS: Em:A0093575"
complement(15734..16180)
/note="match: GSS: Em:AQ142889"
16151..16285
/note="MER46C repeat: matches 181..333 of consensus"
MER46C repeat: matches 181..333 of consensus"
16354..16432
/note="MER46C repeat: matches 2..81 of consensus"
MER46C repeat: matches 2..81 of consensus"
16754..16843
/note="L2 repeat: matches 2641..2750 of consensus"
complement(16918..17359)
/note="match: STS: Em:G23181"
17228..17305
/note="L2 repeat: matches 2668..2746 of consensus"
17564..17668
/note="MER5A repeat: matches 57..161 of consensus"
18161..18307
/note="MIR repeat: matches 6..174 of consensus"
18633..18949

```

Query Matcn 34.7%: Score 35; DB 9; Length 127794;
Best Local Similarity 66.7%: Pred. No. 6.3;
Matches 50; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```

OY 5 cccccctgtatgaggtgtgcttggaatctataggcaaccagagggcgagggcgaggg 64
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 44458 cccccctgtatgaggtgtgcttggaatctataggcaaccagagggcgagggcgaggg 64
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
OY 65 ggaagtcctctggag 79
    |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 44398 gggggcattccggcagc 44384

```

RESULT 7
AF461698
LOCUS AF461698 692 bp mRNA linear VRT 24-JAN-2002

```

DEFINITION  Gallus gallus paired-related homeobox mRNA, partial cds.
ACCESSION  AF461698
VERSION    AF461698.1  GI:18308153
KEYWORDS   chicken.
SOURCE     Gallus gallus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 692)
AUTHORS   Venkataramu,C.R. and Sokolowski,B.H.A.
TITLE     Protein-protein interactions of Kvbeta
JOURNAL   Unpublished
          2 (bases 1 to 692)
AUTHORS   Venkataramu,C.R. and Sokolowski,B.H.A.
TITLE     Direct Submission
JOURNAL   Submitted (21-DEC-2001) Otolaryngology, University of South
          Florida, 12901 Bruce B. Downs Blvd., Tampa, FL 33612, USA

FEATURES
Source     1..692
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            /tissue_type="inner ear"
            /dev_stage="embryonic"
            52..>692
            /codon_start=1
            /product="paired-related homeobox"
            /protein_id="AA167846.1"
            /db_xref="GI:18308154"
            /translation="MASSVAHAMERQALLPARLDGPAGLDNLQAKKNSVSHLLDLE
            AGDMVAAGDEGEGEPGRSLLESGLTSGSDTPQDNDQJNSEKKRKRRNRRTTN
            SSOLALERFERTHYPDARFREDLARVNLTEARVQVWONRAKFRERERMLASK
            NASLKSYSGDYTAVEQPIVPRAPRPTDYLMSGTASPYSAATYSTCTMNSP"

BASE COUNT  158 a      220 c      217 g      97 t

ORIGIN
Query Match      32.9%; Score 33.2; DB 5; Length 692;
Best Local Similarity 67.1%; Pred. No. 59;
Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY  17  ggaatggcctggaatctatggagcaccagagggcgagggagagtcctcctg 76
    |||| |||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   177  GGAGGAGCGGGCCGACATGCTGTGCGCCGCGAGGAGCGCGCGCGCGCGCG

QY  77  gagcctgctg 86
    |||| ||| ||
Db   237  GAGCTTCTTG 246

RESULT  8
LOCUS      S69088
DEFINITION gmxo-x-homeobox [chickens, stage 17 embryo, heart, mRNA, 895 nt].
ACCESSION  S69088
VERSION    S69088.1  GI:545359
KEYWORDS   chicken heart stage 17 embryo.
SOURCE     Gallus gallus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 895)
REFERENCE  Kuratani,S., Matlin,J.F., Wawersik,S., Lilly,B., Elchele,G. and
            Olson,E.N.
TITLE     The expression pattern of the chick homeobox gene gmxoX suggests a
            role in patterning of the limbs and face and in
            compartmentalization of somites
JOURNAL   Dev. Biol. 161 (2), 357-369 (1994)
MEDLINE   94148118
REMARK     GenBank staff at the National Library of Medicine created this
            entry [NCBI gidsq 144347] from the original journal article.
            This sequence comes from Fig. 1A.

```

```

FEATURES
Source     Location/Qualifiers
            1..895
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            1..895
            /gene="gmxoX"
            /note="homeobox"
            31..768
            /gene="gmxoX"
            /note="homeobox protein: This sequence comes from Fig.
            1A"
            1A
            /codon_start=1
            /protein_id="AA29880.1"
            /db_xref="GI:545360"
            /translation="MASSVAHAMERQALLPARLDGPAGLDNLQAKKNSVSHLLDLE
            AGDMVAAGDEGEGEPGRSLLESGLTSGSDTPQDNDQJNSEKKRKRRNRRTTN
            SSOLALERFERTHYPDARFREDLARVNLTEARVQVWONRAKFRERERMLASK
            NASLKSYSGDYTAVEQPIVPRAPRPTDYLMSGTASPYSAATYSTCTMNSPAGM
            NMANSIANLRKLAKETSLORNOVPTYN"

BASE COUNT  254 a      255 c      246 g      140 t

ORIGIN
Query Match      32.9%; Score 33.2; DB 5; Length 895;
Best Local Similarity 67.1%; Pred. No. 55;
Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY  17  ggaatggcctggaatctatggagcaccagagggcgagggagagtcctcctg 76
    |||| |||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   156  GGAGGAGCGGGCCGACATGCTGTGCGCCGCGAGGAGCGCGCGCGCGCGCGCG

QY  77  gagcctgctg 86
    |||| ||| ||
Db   216  GAGCTTCTTG 225

RESULT  9
LOCUS      CHKPRX1
DEFINITION Gallus gallus paired-related homeobox mRNA, complete cds.
ACCESSION  D13433
VERSION    D13433.1  GI:222850
KEYWORDS   paired-related homeobox.
SOURCE     Gallus gallus (Library: lambda gt10) stage 24-26 cDNA to mRNA,
            clones p2 and p7.
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 1084)
REFERENCE  Nohno,T., Koyama,E., Myokai,F., Taniguchi,S., Ohuchi,H., Salto,T.
            and Noji,I.S.
TITLE     A chicken homeobox gene related to Drosophila paired is
            predominantly expressed in the developing limb
JOURNAL   Dev. Biol. 158 (1), 254-264 (1993)
MEDLINE   93321789
COMMENT    Submitted (20-OCT-1992) to DDBJ by:
            Tsutomu Nohno
            Department of Pharmacology
            Kawasaki Medical School
            577 Matsushima
            Kurashiki 701-01
            Japan
            Phone: 086-462-1111
            Fax: 086-462-1199.
            Location/Qualifiers
            1..1084

```

```

gene
  cds
    /organism="Gallus gallus"
    /db_xref="taxon:9031"
    /clone_11b="lambda gt10"
    /dev_stage="Stage 24-26"
    233..967
    /gene="Prx-1"
    233..967
    /gene="Prx-1"
    /codon_start=1
    /product="paired-related homeotic gene product"
    /protein_id="BA02695.1"
    /db_xref="GI:222851"
    /translation="MASSYAHAMERQALLPARLDGPGAGLQIAKKNFVSHLLDLE
AGDWAAGDGGGSGPGLSPLESPTSGSDTPQDDNDOLNSEKKRKRORNTFTN
SSQALALERVEKTHYPDAFVEDLARVNLTEARVQVVFQNRRAKFRNRERMSKN
ASLKSYSQDVTAVEOPTVPRAPRPTDYSWGTASPSYAMATYSTCTCTNAPAGNH
MANSLANLRKAKEVSLQRNVPTVN"
    509..691
    /gene="Prx-1"
    /note="homeobox"

BASE COUNT      258 a      337 c      317 g      172 t
ORIGIN

Query Match      32.9% Score 33.2; DB 5; Length 1084;
Best Local Similarity 67.1%; Pred. No. 53;
Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 17 ggaagtggcctggaaactctataggaccagagggcgagggcgagggagtcctcctg 76
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 GGAGGAGCGCGGCACATGTGCGCGCGAGGAGGAGGCGGCGGCGAGCCGCGCG 417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 77 gagcctggtg 86
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 GAGCTTGTG 427

RESULT 10
LOCUS      DB6982      6335 bp      mRNA      linear      PRI 06-OCT-2001
DEFINITION      Human mRNA for KIAA0229 gene, partial cds.
ACCESSION      DB6982
VERSION      DB6982.1 GI:1504037
KEYWORDS
SOURCE      Homo sapiens male bone marrow myeloblast cell_line:KG-1 cDNA to
      mRNA clone:HA2570.
ORGANISM      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS      Nagase,T., Seki,N., Ishikawa,K., Ohira,M., Kawarabayashi,Y.,
      Ohara,O., Tanaka,A., Kotani,H., Miyajima,N. and Nomura,N.
      Prediction of the coding sequences of unidentified human genes. VI.
      The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
      analysis of cDNA clones from cell line KG-1 and brain
      DNA Res. 3 (5), 321-329 (1996)
      97191544
      2 (bases 1 to 6335)
      Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.
      Direct Submission
      Submitted (02-AUG-1996) Osamu Ohara, Kazusa DNA Research Institute;
      1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
      (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913)
      Location/Qualifiers
      1..6335
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="Chromosome 6. (RH_ID:RH25435)"
      /clone="HA2570"
      /sex="male"
      /cell_line="KG-1"
      /cell_type="myeloblast"
      /tissue_type="bone marrow"
FEATURES
source

```

```

gene
  cds
    1..3544
    /gene="KIAA0229"
    <1..3544
    /gene="KIAA0229"
    /note="similar to human ankyrin 1(S08275)"
    /codon_start=2
    /protein_id="BA013218.1"
    /db_xref="GI:1504038"
    /translation="SMKGRGVSPAGLGLALPGDGRFGSPRLGCSLGEYORVA
LGMKEDELLIARTGHLPAVERKLSKRLISGCGGGGCGGCGGCGGCGGCGGCGG
SSHPVLSLSMMRGPNVNCVDSGYTPPLHAAALNGHDVVEVLLRNDALNVADSKC
YPLHAAKGDADQIVRLLIHQGPSHTFVNQNDNETALNCAAOYCHTEVYKLLLEL
TDPTVRNNKEETPLDLALYGRLEVYMLNANPRLNCSKTKHTPLHLAARGHNAV
VOVLLDGMNSQTEMGSLHAEALGKTDVVOILLAAGTQVINDNHLITLDYR
ELPSKSDQIALILEDHMGKRSKPEVDKPPPOPLISMDISOKSQGDYKATE
LIIIDFQNADEBQTEYLYNAISCHSLDSMASGSSQDDSTNKAEAGVAPAGVPR
BRPPPAKPPPPDEEDHIDTKTFPLTASVLSMRPRLHGSAREDEHRYELLTAE
TKRVLVLDGKTKDHRSSSSSSODSAGODGOVPEQFSLHGSIPVCEYGDPPOLL
CTAGOSHPPDGSPOGACHKASMOLEETGVAPASOPALDQSKRVGYLTGLPTNSR
SHPETLTHRTASPHGAEEDRGARSRAPRTSKPAELKLSRLSKSDSLTCSPT
EDATMGSRSESLNCSIGKKRLKSPSEMEPEIKESIGEGIDFQEROKISGS
RTLBOAGYEWLESIGLOQYRSKLLNFDFVHFLGSMVEEDRLRIGISDPHRRKL
LQAAKSLPKYKALGDGNSPPSPSHDLSGLDDYVHSLSGYSIDYKVNMLEL
VNLKVLVLDGKRRKRIIASLADRYEPPPPPPSOELRCQDLSQSSSLQNDSCG
RSADLLLPDGTGRRRHDSLDPAASRAFRFIQEBHREAKTLPLPSPSLAAYAPQ
SMOQHEKLIFFESGCEANYLGSMILKIDTESTODACAKMKRSTEHMKIPTIILS
ITYKGVFIDASNKNVIAHEIRINSCAADDPDLCTFPAITKDLQSHHYCHVFSY
DVNLTYEILITLGOAFVAYOLALQAKSRATGASAAEMLETSSRPVPRPRVCYKRS
ALPPMDDOAQSHASVWYDPKPRDSKRLSTIN"
    3545..6335
    /gene="Prx-1"
    /note="homeobox"

BASE COUNT      1458 a      1727 c      1766 g      1384 t
ORIGIN

Query Match      32.3% Score 32.6; DB 9; Length 6335;
Best Local Similarity 60.9%; Pred. No. 51;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Oy 4 tccccctgtgtagagctgggaactcttaaggaccaccagagggcgagggcgag 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2918 TCTGCACTAGGGCGCTAGGGGCTGCCAGCCTGGGGGCCCGCAGGCTGACCTTGCCCTCAC 2659
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 64 ggaagctcctcctggagcctggtgcct 90
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2858 GGTCTCTCTCTCTGATCCTGAAGCGCT 2832

RESULT 11
LOCUS      AC091776      54018 bp      DNA      linear      HTG 29-JAN-2002
DEFINITION      Chlamydomonas reinhardtii clone cr-4121, WORKING DRAFT SEQUENCE, 5
      unordered pieces.
ACCESSION      AC091776
VERSION      AC091776.10 GI:18390210
KEYWORDS      HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE      Chlamydomonas reinhardtii.
      Chlamydomonas reinhardtii.
      Chlamydomonas reinhardtii.
      Chlamydomonadaceae; Chlorophyta; Chlorophyceae; Volvocales;
      Chlamydomonadaceae; Chlamydomonas.
ORGANISM
      1 (bases 1 to 54018)
      Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
      Chlamydomonas reinhardtii BAC Clone cr-4121
      Unpublished
      2 (bases 1 to 54018)
      Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
      Direct Submission
      Submitted (02-JUN-2001) Department of Chemistry And Biochemistry,
      The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
      OK 73019, USA
      On Jan 29, 2002 this sequence version replaced gi:16271941.
      ----- Genome Center
      Center: Department of Chemistry And Biochemistry
      The University of Oklahoma

```


Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 5017: contig of 5017 bp in length
5018 5117: gap of unknown length
5118 10018: contig of 4901 bp in length
10019 10118: gap of unknown length
10119 21786: contig of 11668 bp in length
21787 21886: gap of unknown length
21887 33146: contig of 11260 bp in length
33147 33246: gap of unknown length
33247 54018: contig of 20772 bp in length.
Location/Qualifiers

1..54018
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone="cr-4121"
/clone_11b="cr-32m22"

BASE COUNT 9093 a 18192 c 16851 g 9482 t 400 others
ORIGIN

Query Match 32.3%; Score 32.6; DB 2; Length 54018;
Best Local Similarity 63.3%; Pred. No. 32;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 3 gtcctccctgtgagagtgagtggaatctatgagcaccagagggcgaggcgaga 62
DB 53405 GTTGCTGCTGTAGAGGCGGCTGCTGCTTCATGCGACCCGTGGCGCGCGCGC 53346
OY 63 gggagagtcctctgagcc 81
DB 53345 GGTGGCCCGCGCCTGGGCC 53327

RESULT 12
AC090435 65898 bp DNA linear HTG 29-JAN-2002
LOCUS Chlamydomonas reinhardtii clone cr-32m22, WORKING DRAFT SEQUENCE, 4
DEFINITION unordered pieces.
AC090435
VERSION AC090435.16 GI:18390207
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Chlamydomonadaceae: Chlamydomonas.
AUTHORS Eukaryota: Viridiplantae: Chlorophyta: Chlorophyceae: Volvocales;
TITLE Chlamydomonas reinhardtii BAC Clone cr-32m22
JOURNAL (bases 1 to 65898)
UNPUBLISHED Wu, H., Lin, S., Jia, H., Dutcher, S. and Roe, B.A.
Chlamydomonas reinhardtii

REFERENCE
AUTHORS Wu, H., Lin, S., Jia, H., Dutcher, S. and Roe, B.A.
TITLE Chlamydomonas reinhardtii BAC Clone cr-32m22
JOURNAL (bases 1 to 65898)
AUTHORS Wu, H., Lin, S., Jia, H., Dutcher, S. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Jan 29, 2002 this sequence version replaced g1:18139395.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 4454: contig of 4454 bp in length
4455 4554: gap of unknown length
4555 18724: contig of 14170 bp in length
18725 18824: gap of unknown length
18825 30657: contig of 11833 bp in length
30658 30757: gap of unknown length
30758 65898: contig of 35141 bp in length.
Location/Qualifiers

1..65898
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone="cr-32m22"
/clone_11b="cr-32m22"

BASE COUNT 10668 a 22600 c 21674 g 10642 t 314 others
ORIGIN

Query Match 32.3%; Score 32.6; DB 2; Length 65898;
Best Local Similarity 63.3%; Pred. No. 30;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 3 gtcctccctgtgagagtgagtggaatctatgagcaccagagggcgaggcgaga 62
DB 30258 GTTGCTGCTGTAGAGGCGGCTGCTGCTTCATGCGACCCGTGGCGCGCGCGC 30199
OY 63 gggagagtcctctgagcc 81
DB 30198 GGTGGCCCGCGCCTGGGCC 30180

RESULT 13
AL133352 16199 bp DNA linear HTG 04-APR-2000
LOCUS Homo sapiens chromosome 10 clone RP11-411b6, ** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
AC090435
VERSION AL133352.10 GI:7329506
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
TITLE (bases 1 to 16199)
JOURNAL Direct Submission
Submitted (04-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Mar 26, 2000 this sequence version replaced g1:7327699.
IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Contig_ID: 00484 Length: 1515bp
Contig_ID: 00808 Length: 1515bp
Contig_ID: 01411 Length: 1060bp.

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 157824: contig of 157824 bp in length
* 157825 158624: gap of 800 bp

72056 72155: gap of unknown length
72156 75560: contig of 3405 bp in length
75561 75660: gap of unknown length
75661 79220: contig of 3560 bp in length
79221 79320: gap of unknown length
79321 81893: contig of 2573 bp in length
81894 81994: gap of unknown length
81994 84523: contig of 2530 bp in length
84524 84623: gap of unknown length
84624 87701: contig of 3078 bp in length
87702 87801: gap of unknown length
87802 89855: contig of 2054 bp in length
89856 89955: gap of unknown length
89956 92884: contig of 2929 bp in length
92885 92984: gap of unknown length
92985 94489: contig of 1505 bp in length
94490 94589: gap of unknown length
94590 97221: contig of 2632 bp in length
97222 97321: gap of unknown length
97322 99641: contig of 2320 bp in length
99642 99741: gap of unknown length
99742 101792: contig of 2051 bp in length
101793 101892: gap of unknown length
101893 104245: contig of 2353 bp in length
104246 104345: gap of unknown length
104346 106914: contig of 2569 bp in length
106915 107014: gap of unknown length
107015 108995: contig of 1981 bp in length
108996 109095: gap of unknown length
109096 110468: contig of 1373 bp in length
110469 110568: gap of unknown length
110569 113588: contig of 3020 bp in length
113589 113688: gap of unknown length
113689 116211: contig of 2523 bp in length
116212 116311: gap of unknown length
116312 116773: contig of 2362 bp in length
116774 118774: gap of unknown length
118774 120028: contig of 1255 bp in length
120029 120128: gap of unknown length
120129 122287: contig of 2159 bp in length
122288 122387: gap of unknown length
122389 125021: contig of 2634 bp in length
125022 125121: gap of unknown length
125122 126584: contig of 1463 bp in length
126585 126684: gap of unknown length
126685 128864: contig of 2180 bp in length
128865 128964: gap of unknown length
128965 130025: contig of 1061 bp in length
130026 130125: gap of unknown length
130126 131491: contig of 1366 bp in length
131492 131591: gap of unknown length
131592 133067: contig of 1476 bp in length
133068 133167: gap of unknown length
133168 134803: contig of 1636 bp in length
134804 134903: gap of unknown length
134904 136204: contig of 1301 bp in length
136205 136304: gap of unknown length
136305 137709: contig of 1405 bp in length
137710 137809: gap of unknown length
137810 139029: contig of 1220 bp in length
139030 139130: gap of unknown length
139130 140283: contig of 1134 bp in length
140284 140383: gap of unknown length
140384 141708: contig of 1325 bp in length
141709 141808: gap of unknown length
141809 142979: contig of 1171 bp in length
142980 143079: gap of unknown length
143080 144210: contig of 1131 bp in length
144211 144310: gap of unknown length
144311 145426: contig of 1116 bp in length
145427 145526: gap of unknown length
145527 147619: contig of 2093 bp in length
147620 147719: gap of unknown length

Query Match
Best Local Similarity 32.3%; Score 32.6; DB 2; Length 172748;
Pred. No. 25;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Oy 4 tccccccgagagagatggagctggaatctatgagcaccagagggcgagagcgag 63
Db 158714 TCcCGcAGtGAGcCGcCGcGTGAGTGTGTGcAGCTCCGCGGTGGCGGAG 158655
Oy 64 gggagtcctcctgagcctggtgcctagaagccc 98
Db 158654 GCGCGTCTTCTTACGCGCAGCAGCATACMAAGCTC 158620
RESULT 15
AC008744/c 203200 bp DNA linear PRI 22-JAN-2001
LOCUS AC008744
DEFINITION Homo sapiens chromosome 19 clone CTD-2561020, complete sequence.
AC008744
VERSION AC008744.6 GI:12331466
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 203200)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE
AUTHORS 2 (bases 1 to 203200)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE
AUTHORS 3 (bases 1 to 203200)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
COMMENT Submitted (22-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 22, 2001 this sequence version replaced gi:11178043.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.8.
Note:
Clipped left at overlap with AC008738.
Clipped right at overlap with AC010485.
STS Content:
SHGC-35371 G28590
WI-6761 G06103
WI-7774 G06697
SHGC-33750 G29472.
Location/Qualifiers
1..203200
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2561020"
BASE COUNT 52932 a 51214 c 50304 g 48750 t
ORIGIN

```

Query Match          32.3%; Score 32.6; DB 9; Length 203200;
Best Local Similarity 58.9%; Pred. No. 24;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 7 cccctgggagagagtggtctgggaatctatbaggcacccagagggcggggagggaggg 66
      ||||| | ||||| ||| | | | | | | | | | | | | | | | | | | | | |
Db 133557 cccctgcaggggagtggtctgggaatctatbaggcacccagagggcggggagggaggg 66
      ||||| | ||||| ||| | | | | | | | | | | | | | | | | | | | | |

QY 67 agtctctctggagccttggtgcccctagaagccacg 101
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133497 atagcccatgcagcccttcagccacg 133463
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: October 3, 2002, 14:46:57
 Job time: 11937 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:16 ; Search time 719.93 Seconds
(without alignments)
240.866 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700

Perfect score: 1 aggtccccctggtgaggaag.....tggtgccccagaagccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N.Geneseq_032802:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1722	22	AAH43683
2	32.6	32.3	550	22	ABA61253
3	32.6	32.3	550	22	ABA29092
4	32.6	32.3	550	22	AAK09549
5	32.6	32.3	550	22	AAK35441
6	32.6	32.3	550	22	AAI17045
7	32.6	32.3	550	22	AAI14154
8	32.6	32.3	680	22	AAK76637
9	32.6	32.3	3544	23	AAH81565

C	10	32.4	32.1	3750	23	AAH74057	DNA encoding novel
C	11	32.4	32.1	3889	22	AAH90107	Human bone marrow
C	12	32.4	32.1	3892	22	AAH90056	Human bone marrow
C	13	32.4	32.1	4886	22	AAH89943	Human bone marrow
C	14	32.2	31.9	231	22	ABA73753	Human foetal liver
C	15	32.2	31.9	231	22	ABA38943	Probe #17409 for g
C	16	32.2	31.9	231	22	AAK22202	Human brain expres
C	17	32.2	31.9	231	22	AAK48368	Human bone marrow
C	18	32.2	31.9	231	22	AAI26223	Probe #16156 for g
C	19	32.2	31.9	231	22	AAI54194	Probe #22880 used
C	20	31.4	31.1	17761	22	AAK79415	Human immune/haema
C	21	31	30.7	686	22	AAH08414	Human cDNA clone (
C	22	31	30.7	1030	21	AAK77940	Human cancer assoc
C	23	31	30.7	1105	22	AAH77880	Human ATRX coding
C	24	31	30.7	1374	22	AAH18673	Human cDNA sequenc
C	25	30.6	30.3	412	22	AAK59971	Human immune/haema
C	26	29.8	29.5	1688	22	AAH34940	Human colon cancer
C	27	29.4	29.1	1951	22	ABA44228	Human breast cell
C	28	29.4	29.1	1951	22	ABA54678	Human foetal liver
C	29	29.4	29.1	1951	22	ABA24460	Probe #2926 for ge
C	30	29.4	29.1	1951	22	AAK02965	Human brain expres
C	31	29.4	29.1	1951	22	AAK28411	Human bone marrow
C	32	29.4	29.1	1951	22	AAI12976	Probe #2909 for ge
C	33	29.4	29.1	1951	22	AAI13435	Probe #3021 used t
C	34	29.4	29.1	1951	22	AAI02894	Probe #2885 used t
C	35	28.8	28.5	17084	22	AAK70764	Human immune/haema
C	36	28.8	28.5	17084	22	AAK71755	Human immune/haema
C	37	28.6	28.3	19062	21	AAH21280	Human low adenosin
C	38	28.6	28.3	23142	21	AAA35158	Human adenosine re
C	39	28.6	28.3	30417	21	AAH21282	Human low adenosin
C	40	28.6	28.3	30417	21	AAA35160	Human adenosine re
C	41	28.4	28.1	96774	22	AAI02893	Human reproductive
C	42	28.4	28.1	20387	19	AAV62135	HSV-2 strain SB5 C
C	43	28.4	28.1	26338	19	AAV62134	HSV-2 strain SB5 C
C	44	28.4	28.1	117213	19	AAV62176	HSV-2 strain SB5 C
C	45	28.4	28.1	154746	24	AAH25519	Human herpesvirus

ALIGNMENTS

RESULT 1	AAH43683	AAH43683 standard; DNA; 1722 BP.
ID	AAH43683	
XX	AAH43683:	
XX	21-JAN-2002 (first entry)	
DT	PRKAG3 intron 4 - intron 10.	
DE		
XX	Human: AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;	
KW	metabolic disease; diabetes; obesity; substitution; ds.	
KX	Homo sapiens.	
OS		
XX		
XX		
FT	Key	Location/Qualifiers
FT	intron	1..13
FT		/tag= a
FT		/number= "Intron 4"
FT	exon	/note= "3' portion of Intron 4"
FT		14..95
FT		/tag= b
FT		/number= "Exon 5"
FT	intron	96..552
FT		/tag= c
FT		/number= "Intron 5"
FT	exon	553..611
FT		/tag= d
FT		/number= "Exon 6"
FT	intron	612..736
FT		/tag= e
FT		/number= "Intron 6"


```
XX 23-JAN-2002 (first entry)
XX
XX Probe #7558 for gene expression analysis in human heart cell sample.
DE
XX
XX Human: gene expression: heart; microarray: vascular system; probe:
KM cardiovascular disease: hypertension; cardiac arrhythmia;
KM congenital heart disease: ss.
XX
XX Homo sapiens.
OS
XX WO200157274-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00666.
PR
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
PT
XX
XX Claim 1: SEQ ID No 7558; 530bp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other:
SQ
XX
XX Query Match 32.3%; Score 32.6; DB 22; Length 550;
XX Best Local Similarity 60.9%; Pred. No. 4.3;
XX Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
XX
XX 10 ctggtgagagtggtggaattctatggaaccacagagggcgaggcgaggaagt 69
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 ctgattgaagactagcgcgaggcgacagctgtgacacagatggcgagggtgctgggag 244
XX
XX 70 cctctgagagcctgtgtccctagaagc 96
XX || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 ccaaaagtaaccagcttccctccaac 271
XX
XX
XX RESULT 4
XX ID AAK09549 standard; DNA: 550 BP.
XX AC AAK09549;
XX XX
XX DT 05-NOV-2001 (first entry)
XX XX
```

```
DE Human brain expressed single exon probe SEQ ID NO: 9540.
XX
XX Human: brain expressed exon; gene expression analysis: probe:
KM microarray: Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer; ss.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PR
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
PT
XX
XX Example 4; SEQ ID NO: 9540; 650bp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
XX Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other:
SQ
XX
XX Query Match 32.3%; Score 32.6; DB 22; Length 550;
XX Best Local Similarity 60.9%; Pred. No. 4.3;
XX Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
XX
XX 10 ctggtgagagtggtggaattctatggaaccacagagggcgaggcgaggaagt 69
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 ctgattgaagactagcgcgaggcgacagctgtgacacagatggcgagggtgctgggag 244
XX
XX 70 cctctgagagcctgtgtccctagaagc 96
XX || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 ccaaaagtaaccagcttccctccaac 271
XX
XX
XX RESULT 5
XX ID AAK35441 standard; DNA: 550 BP.
XX AC AAK35441;
XX XX
XX DT 06-NOV-2001 (first entry)
XX XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 9998.
XX
XX Human: bone marrow expressed exon; gene expression analysis: probe:
KM microarray: cancer; leukemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
OS
XX WO200157276-A2.
XX XX
```

```
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4: SEQ ID NO: 9998; 658bp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 550 BP: 154 A; 139 C; 127 G; 130 T; 0 other;
XX
XX Query Match 32.3%; Score 32.6; DB 22; Length 550;
XX Best Local Similarity 60.9%; Pred. No. 4.3;
XX Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
XX
XX Oy 10 ctggtgaggagtggtggtgaattctatggcaccagaaggcgagggaggt 69
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 185 ctgattgaagactaaggcgaggcgtgtggtcaccagatggcagggtctctggggaag 244
XX
XX Oy 70 cctctgagagcctgtgctccctagaagc 96
XX || | | | | | | | | | | | | | |
XX Db 245 ccaaaagtaaccagcttcctccaac 271
XX
XX RESULT 6
XX AAI17045
XX ID AAI17045 standard; DNA: 550 BP.
XX
XX AC AAI17045;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Probe #6978 for gene expression analysis in human cervical cell sample.
XX
XX KW Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX OS Homo sapiens.
XX
XX PN MO200157278-A2.
XX
XX PD 09-AUG-2001.
XX
XX PE 30-JAN-2001; 2001WO-US00670.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
```

```
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25: SEQ ID NO 6978; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 550 BP: 154 A; 139 C; 127 G; 130 T; 0 other;
XX
XX Query Match 32.3%; Score 32.6; DB 22; Length 550;
XX Best Local Similarity 60.9%; Pred. No. 4.3;
XX Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
XX
XX Oy 10 ctggtgaggagtggtggtgaattctatggcaccagaaggcgagggaggt 69
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 185 ctgattgaagactaaggcgaggcgtgtggtcaccagatggcagggtctctggggaag 244
XX
XX Oy 70 cctctgagagcctgtgctccctagaagc 96
XX || | | | | | | | | | | | | | |
XX Db 245 ccaaaagtaaccagcttcctccaac 271
XX
XX RESULT 7
XX AAI41154
XX ID AAI41154 standard; DNA: 550 BP.
XX
XX AC AAI41154;
XX
XX DT 17-OCT-2001 (first entry)
XX
XX DE Probe #9840 used to measure gene expression in human placenta sample.
XX
XX KW Probe: microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN MO200157272-A2.
XX
XX PD 09-AUG-2001.
XX
XX PE 30-JAN-2001; 2001WO-US00663.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
```


XX WP1: 2001-488897/53.
DK
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
PT
XX
PS Claim 25: SEQ ID NO 9840: 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 550 BP: 154 A; 139 C; 127 G; 130 T; 0 other;

Query Match	32.3%	Score 32.6	DB 22	Length 550
Best Local Similarity	60.98%	Pred. No. 4.3		
Matches 53: Conservative	0	Mismatches 34	Indels 0	Gaps 0

QY 10 ctggtcaggaagtcggctctcggaatcttatcggccaccagaagggcgggctcgagggagtc 69
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 185 ctgatgaagactaagcgcgggcgaagcctcttgcacccaagatcggccagggctctctgggaag 244

Oy	70	cctcctgagcctgltgccctagaac	96
	11	1111111111111111	
Db	245	ccaaagtaaccagcttcctccaac	271

RESULT	8
AAK76637	
ID	AAK76637 standard; DNA; 680 BP.

AC	AAK76637;
XX	
DT	07-NOV-2001 (first entry)

DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31449.
DE	

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KM cytostatic; gene therapy; vaccine; metastasis; ds.

05 Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 24-FEB-2000; 2000US-0184664.

PR 16-MAR-2000; 2000US-0189874.

PR 18-APR-2000; 2000US-0198123.

PR 07-JUN-2000; 2000US-0209467.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217407

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218200

PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964

PR 14-AUG-2000; 2000US-0224518.
PB 14-AUG-2000; 2000US-0224518

PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214

PR 14-AUG-2000; 2000US-0225214.

PR	14 - AUG -2000	2000US -02255266
PR	14 - AUG -2000	2000US -02252687
PR	14 - AUG -2000	2000US -02252658
PR	14 - AUG -2000	2000US -02252547
PR	14 - AUG -2000	2000US -02255757
PR	14 - AUG -2000	2000US -02255758
PR	14 - AUG -2000	2000US -02255759
PR	18 - AUG -2000	2000US -02262679
PR	22 - AUG -2000	2000US -02266681
PR	22 - AUG -2000	2000US -02266688
PR	22 - AUG -2000	2000US -02271892
PR	23 - AUG -2000	2000US -02282924
PR	30 - AUG -2000	2000US -02289247
PR	01 - SEP -2000	2000US -02292827
PR	01 - SEP -2000	2000US -02293433
PR	01 - SEP -2000	2000US -02293434
PR	01 - SEP -2000	2000US -02293435
PR	03 - SEP -2000	2000US -02295019
PR	05 - SEP -2000	2000US -02295113
PR	06 - SEP -2000	2000US -02304337
PR	06 - SEP -2000	2000US -02304338
PR	08 - SEP -2000	2000US -02312422
PR	08 - SEP -2000	2000US -02312423
PR	08 - SEP -2000	2000US -02312424
PR	08 - SEP -2000	2000US -02314133
PR	08 - SEP -2000	2000US -02314134
PR	08 - SEP -2000	2000US -02312060
PR	08 - SEP -2000	2000US -02312061
PR	12 - SEP -2000	2000US -02311968
PR	14 - SEP -2000	2000US -02323397
PR	14 - SEP -2000	2000US -02323398
PR	14 - SEP -2000	2000US -02323399
PR	14 - SEP -2000	2000US -02324000
PR	14 - SEP -2000	2000US -02324001
PR	14 - SEP -2000	2000US -02330663
PR	14 - SEP -2000	2000US -02354884
PR	14 - SEP -2000	2000US -02330655
PR	21 - SEP -2000	2000US -02358327
PR	21 - SEP -2000	2000US -02342424
PR	23 - SEP -2000	2000US -02349977
PR	23 - SEP -2000	2000US -02349988
PR	25 - SEP -2000	2000US -02354884
PR	26 - SEP -2000	2000US -02354884
PR	27 - SEP -2000	2000US -02358327
PR	28 - SEP -2000	2000US -02363527
PR	29 - SEP -2000	2000US -02363527
PR	29 - SEP -2000	2000US -02363568
PR	29 - SEP -2000	2000US -02363569
PR	29 - SEP -2000	2000US -02363602
PR	02 - OCT -2000	2000US -02370337
PR	02 - OCT -2000	2000US -02370338
PR	02 - OCT -2000	2000US -02370339
PR	02 - OCT -2000	2000US -02370345
PR	13 - OCT -2000	2000US -02393935
PR	13 - OCT -2000	2000US -02393937
PR	20 - OCT -2000	2000US -02418089
PR	20 - OCT -2000	2000US -02418090
PR	20 - OCT -2000	2000US -02418109
PR	20 - OCT -2000	2000US -02418116
PR	01 - NOV -2000	2000US -02446217
PR	08 - NOV -2000	2000US -02466754
PR	08 - NOV -2000	2000US -02466755
PR	08 - NOV -2000	2000US -02466776
PR	08 - NOV -2000	2000US -02466777
PR	08 - NOV -2000	2000US -02466778
PR	08 - NOV -2000	2000US -02465224

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246538.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Barash SC, Ruben SM;
P1 WPI: 2001-483426/52.
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX

PS Disclosure: SEQ ID NO 31449; 3071bp + Sequence Listing; English.
XX
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (II)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients' own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (II) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX

XQ Sequence 680 BP; 116 A; 213 C; 185 G; 166 T; 0 other:
XX

Query Match	32.3%	Score 32.6	DB 22	Length 680
Best Local Similarity	58.9%	Pred. No. 4.3		
Matches 56	Conservative 0	Mismatches 39	Indels 0	Gaps 0
Db	629	atagccatcgacgcatgacgcttgcagccagc	663	
Qy	67	agtcctccgagcctggtgcctctagaagccagc	101	
Db	569	ccccctgtagagagctggagcctctatagcaccacagagggcgagggcgagggg	66	
Qy	67	ccccctgtagagagctggagcctctatagcaccacagagggcgagggcgagggg	66	
Db	629	atagccatcgacgcatgacgcttgcagccagc	663	
RESULT 9				
AAS81565/C				
ID	AAS81565	standard	CDNA: 3544	BP.
XX	AAS81565			
XX	13-FEB-2002	(first entry)		
XX	DNA encoding novel human diagnostic protein #17369.			
DE	Human: chromosome mapping; gene mapping; gene therapy; forensic:			
XX	food supplement; medical imaging; diagnostic; genetic disorder; ss.			
XX	Homo sapiens.			
OS	WO200175067-A2.			
PN	11-OCT-2001.			
XX	30-MAR-2001; 2001WO-US08631.			
XX	31-MAR-2000; 2000US-0540217.			
PR	23-AUG-2000; 2000US-0649167.			
PA	(HYSE-) HYSEQ INC.			
PI	Drmanac RT, Liu C, Tang YT;			
PT	WPI: 2001-639362/73.			
DR	P-PSDB; ABG17378.			
XX	New isolated polynucleotide and encoded polypeptides, useful in			
PT	diagnostics, forensics, gene mapping, identification of mutations			
PT	responsible for genetic disorders or other traits and to assess			
PT	biodiversity			
XX	Claim 1; SEQ ID No 17369; 103bp; English.			
XX	The invention relates to isolated polynucleotide (I) and			
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,			
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome			
CC	and gene mapping, and in recombinant production of (II). The			
CC	polynucleotides are also used in diagnostics as expressed sequence tags			
CC	for identifying expressed genes. (I) is useful in gene therapy techniques			
CC	to restore normal activity of (II) or to treat disease states involving			
CC	(II). (II) is useful for generating antibodies against it, detecting or			
CC	quantitating a polypeptide in tissue, as molecular weight markers and as			
CC	a food supplement. (II) and its binding partners are useful in medical			
CC	imaging of tissues expressing (II). (I) and (II) are useful for treating			
CC	disorders involving aberrant protein expression or biological activity.			
CC	The polypeptide and polynucleotide sequences have applications in			
CC	diagnostics, forensics, gene mapping, identification of mutations			
CC	responsible for genetic disorders or other traits to assess biodiversity			
CC	and to produce other types of data and products dependent on DNA and			
CC	amino acid sequences. AAS64197-AAS94564 represent novel human			
CC	diagnostic coding sequences of the invention.			
CC	Note: The sequence data for this patent did not appear in the printed			

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3544 BP; 848 A; 1024 C; 1040 G; 632 T; 0 other;

Query Match 32.3%; Score 32.6; DB 23; Length 3544;
Best Local Similarity 60.9%; Pred. No. 4.7;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 4 tccccctgtgagagtgagctgggaatctatgagcaccagagggcgagag 63
DB 2918 TCTGACACTGGCGGCTAGGGGCTCCAGGCTTGCGCGCCGAGGCTACGCTTGCCCTCAC 2859

OY 64 gggagtcctctgagagcctgtgccct 90
DB 2858 GGTGCTCTCTCTGATCTGATCGGCT 2832

RESULT 10
AA574057/C
ID AA574057 standard; cDNA: 3750 BP.

XX AA574057;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #9861.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001MO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PsDB: ABC09870.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

PS Claim 1: SEQ ID No 9861; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AA564197-AA594364 represent novel human

CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 3750 BP; 841 A; 1043 C; 966 G; 900 T; 0 other;

Query Match 32.1%; Score 32.4; DB 23; Length 3750;
Best Local Similarity 60.0%; Pred. No. 5.4;
Matches 54; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 10 ctgttgagagtgagctgggaatctatgagcaccagagggcgagagag 69
DB 475 CTGATGAAGACTAGGCGGCAAGGCTGTGGCAGCAGATGGCAGGGCTGTGGGAGG 416

OY 70 ccttcctgagagcctggtgacctagaagccca 99
DB 415 CCAAAAGTAACCAAGCTTCCCTCCAAAGCA 386

RESULT 11
AAH90107/C
ID AAH90107 standard; cDNA: 3889 BP.

XX AAH90107;

XX 01-OCT-2001 (first entry)

DE Human bone marrow cDNA, SEQ ID NO: 464.

XX Human: bone marrow; antiinflammatory; cytostatic; neuroprotective;

XX antiviral; antibacterial; antifungal; anti-HIV; haemostatic;

XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; infection; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.

OS Homo sapiens.

PN WO200153453-A2.

PD 26-JUL-2001.

PF 23-DEC-2000; 2000MO-US34960.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 30-NOV-2000; 2000US-0250583.

PA (HYSE-) HYSEQ INC.

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;

PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Drmanac RT;

DR WPI: 2001-488707/53.

DR P-PsDB: AAM00988.

XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX Claim 1: Page 574-579; 648pp; English.
XX The present sequence is one of 251 novel human polynucleotides
XX expressed in the bone marrow. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal

PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1: Page 259-260; 648bp; English.
XX
CC The present sequence is one of 251 novel human polynucleotides
CC expressed in the bone marrow. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. hemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
XX Sequence 4886 BP; 1117 A; 1370 C; 1266 G; 1133 T; 0 other:
SO
Query Match 32.1%; Score 32.4; DB 22; Length 4886;
Best Local Similarity 60.0%; Pred. No. 5.5;
Matches 54; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 10 ctggtagagagtgaggctgggaatcttatggcaccagagggcgaggagggagt 69
DB 1616 CTCATCTAACACTAGCGCGCGCAAGCCTCTGCACACAGATGGCAGGGCTCTGGGAGC 1557
QY 70 cctcctggagcctgtgtgccttagagccca 99
DB 1556 CCAAAAGTAACCACTTCCCTCCAAAGCA 1527
RESULT 14
ABA73753
ID ABA73753 standard; DNA: 231 BP.
XX
XX ABA73753;
AC
XX 01-FEB-2002 (first entry)
DT
XX Human foetal liver single exon nucleic acid probe #22058.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KM
XX Homo sapiens.
OS
XX WO200157277-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00669.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PR
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI: 2001-483447/52.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human foetal liver -
XX

PS Claim 4: SEQ ID NO 22058; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 231 BP; 54 A; 57 C; 63 G; 57 T; 0 other:
SO
Query Match 31.9%; Score 32.2; DB 22; Length 231;
Best Local Similarity 61.2%; Pred. No. 5.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 10 ctggtgagagtgaggctgggaatcttatggcaccagagggcgaggagggagt 69
DB 146 ctgattgaagactagagcgaggctgtgacacagatggcgagggtctg99a9g 205
QY 70 cctcctggagcctgtgtgccttagaa 94
DB 206 ccaaaagtaaccagcttccctccaa 230
RESULT 15
ABA38943
ID ABA38943 standard; DNA: 231 BP.
XX
XX ABA38943;
AC
XX 23-JAN-2002 (first entry)
DT
XX Probe #17409 for gene expression analysis in human heart cell sample.
DE
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease; ss.
KM
XX Homo sapiens.
OS
XX WO200157274-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00666.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PR
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI: 2001-488899/53.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 4; SEQ ID NO 17409; 530pp; English.
PS
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC

CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging, and
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

XX
 SQ Sequence 231 BP: 54 A: 57 C: 63 G: 57 T: 0 other:

Query Match 31.9%: Score 32.2: DB 22: Length 231;
 Best Local Similarity 61.2%: Pred. No. 5.3:
 Matches 52: Conservative 0: Mismatches 33: Indels 0: Gaps 0:

OY 10 ctggctgagagctggcgtggaatctatgagcaccagagggcgagggaggt 69
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 146 ctgatatgaactagcgagggcgagggcgtgtgacacagatggcaggggtgctggggagg 205
 OY 70 cctcctgagggcctggtgacagaa 94
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 206 ccaaaagtaacacagcttccctccaa 230

Search completed: October 3, 2002, 16:30:18
 Job time: 14288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:14:46 ; Search time 5701.1 Seconds

(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700

Perfect score: 101

Sequence: 1 agtgcctccctggtgagagag.....tgggtccctagaagccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcp1:*
7: em_estro:*
8: em_hic:*
9: qb_est1:*
10: qb_est2:*
11: qb_hic:*
12: qb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	35.6	682	10	B1960023
2	35.2	34.9	833	10	BE559614
3	35	34.7	377	9	AW15288
4	33	32.7	872	9	AL530561
5	32.8	32.5	656	10	B1960536
6	32.6	32.3	282	9	AA872225
7	32.6	32.3	354	9	AA026388
8	32.6	32.3	539	10	BF808916
9	32.2	31.9	685	9	AL585199
10	32	31.7	735	10	BG309093
11	31.8	31.5	382	10	BE600720
12	31.8	31.5	463	10	BE600139
13	31.8	31.5	474	10	BE598197
14	31.8	31.5	487	10	BE600794
15	31.6	31.3	695	12	AG083214
16	31.6	31.3	889	10	B1951430
17	31.6	31.3	889	10	B1951430

C 18	31.4	31.1	452	9	AI863974	AI863974 w154c05. x
C 19	31.2	30.9	407	9	AU162756	AU162756 AU162756
C 20	31.2	30.9	516	9	AU162755	AU162755 AU162755
C 21	31.2	30.9	539	9	AL514039	AL514039 AL514039
C 22	31.2	30.9	688	9	AL514251	AL514251 AL514251
C 23	31.2	30.9	883	12	A2527770	A2527770 ENTDA48TF
C 24	31	30.7	242	10	BF685935	BF685935 602140240
C 25	31	30.7	469	9	AI087063	AI087063 0Y70609. x
C 26	31	30.7	526	10	BE906990	BE906990 6021498803
C 27	31	30.7	529	10	BC481234	BC481234 602258560
C 28	31	30.7	531	9	BE019039	BE019039 BD46b1. y
C 29	31	30.7	535	10	BE250790	BE250790 601106982
C 30	31	30.7	535	10	BE296883	BE296883 601176963
C 31	31	30.7	537	10	BE294571	BE294571 601176144
C 32	31	30.7	538	10	BE298630	BE298630 601119416
C 33	31	30.7	542	10	BE295054	BE295054 601173761
C 34	31	30.7	553	10	B1160864	B1160864 602864791
C 35	31	30.7	557	10	B1767103	B1767103 603054158
C 36	31	30.7	572	10	BE783728	BE783728 601471127
C 37	31	30.7	581	9	AL538073	AL538073 AL538073
C 38	31	30.7	606	9	AW663549	AW663549 EST375622
C 39	31	30.7	610	10	BE300397	BE300397 6009444511
C 40	31	30.7	611	10	BT254382	BT254382 602876691
C 41	31	30.7	617	10	BE271259	BE271259 600943675
C 42	31	30.7	629	10	BF794131	BF794131 602255503
C 43	31	30.7	640	9	AL531142	AL531142 AL531142
C 44	31	30.7	649	9	AL531376	AL531376 AL531376
C 45	31	30.7	658	10	BF982575	BF982575 602305702

ALIGNMENTS

RESULT 1
LOCUS B1960023
DEFINITION HVSMEN0022008f Hordeum vulgare rachis EST library HVCDNA0015
(normal) Hordeum vulgare cDNA clone HVSMEN0022008f, mRNA sequence.
ACCESSION B1960023
VERSION B1960023.1 GI:16311278
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae ; Triticeae; Hordeum.
REFERENCE Wing, R., Close, T.J., Kleinof, A., Wise, R., Chin, A., Begum, D., Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Gates, R. and Main, D.
Development of a genetically and physically anchored EST resource for barley genomics: Morex rachis cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 121
Seq primer: AATTACCTCCTCAAGG
High quality sequence start: 2
High quality sequence stop: 553.
Location/Qualifiers
1..682
/organism="Hordeum vulgare"
/cultivar="Morex"
/cd_xref="taxon:4513"
/clone="HVSMEN0022008f"
/clone_id="Hordeum vulgare rachis EST library HVCDNA0015 (normal)"
/tissue_type="Rachis"

FEATURES

source

```

/lab_host="TJc121"
/Note-Vector: pBluescript SK(-): Site_1: EcoRI: Site_2:
XhoI: Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachises were excised and
frozen in liquid nitrogen (Kleinholz lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared, poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
see Close TJ, Wing R, Kleinholz A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/g99pages/bgn/31/cover.html)"
BASE COUNT      75 a      262 c      174 g      171 t
ORIGIN

```

```

Query Match      35.6% Score 36: DB 10: Length 682;
Best Local Similarity 60.0% Pred. No. 47;
Matches 60: Conservative 0: Mismatches 40: Indels 0: Gaps 0:

```

```

OY 1 aggtcccccgtgtagagctggtggaatctatgagaccagagggcgagggcg 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 ACGGCCCCCGCGTGGAGAGGGCGCGCAAGCTGATCTGCACAGATGCGCGACG 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 gaggggagatcctcctgagagcctggtgcccctagaagccac 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 ACGCGTCCCGACTCGCGCGCGCTGCTGCGACCTCTC 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
LOCUS BE559614 833 bp mRNA linear EST 15-AUG-2000
DEFINITION BE559614 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3688176 5',
mRNA sequence.
ACCESSION BE559614
VERSION BE559614.1 GI:9803323
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-femail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM380 row: p column: 01
High quality sequence stop: 687.
Location/Qualifiers
1. 833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3688176"
/clone_lib="NIH_MGC_8"
FEATURES
SOURCE

```

```

/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/Note-Organ: Lymph: Vector: pORF7: Site_1: XhoI: Site_2:
EcoRI: cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      215 a      175 c      257 g      186 t
ORIGIN

```

```

Query Match      34.9% Score 35.2: DB 10: Length 833;
Best Local Similarity 65.0% Pred. No. 74;
Matches 52: Conservative 0: Mismatches 28: Indels 0: Gaps 0:

```

```

OY 7 cccttgtagagagctggagatctatgagaccagagggcgagggcg 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 739 CCCCTGCTTGATGAGTGGCGGTGCGATGCGACCTCGCGCGCGCGCTGGCG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 67 agtcctctgagagcctgtg 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 799 AAGCCCTGGGCGCGCTGTG 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
LOCUS AM315288 377 bp mRNA linear EST 25-APR-2001
DEFINITION AM315288 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AM315288
VERSION AM315288.1 GI:6744544
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

```

```

REFERENCE Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
1 (bases 1 to 377)
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perta,G., Holt,I., Karaymicheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4356
Fax: 402 762 4390
Email: smitht@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the "minscore 20
and mismatch 12 options.
PCR primers
FORWARD: AGGAACACGATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 142 row: D column: 8
Seq primer: ATTACGTGACACTATAG.
Location/Qualifiers
1. 377
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/Note-Vector: PCMV SPORT6; Site_1: XhoI; Site_2: XhoI;
library made from pooled tissue from testis, thymus,
semitendonsus muscle, longissimus muscle, pancreas,

```

```

FEATURES
SOURCE

```


REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 539) Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,I.A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.-J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE COMMENT	Contact: Stimpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL: http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-CI0154-t- 151100-001-h12&t3=2000-11-15&t4=1)
FEATURES source	Seq primer: puc 18 forward High quality sequence start: 20 High quality sequence stop: 539. Location/Qualifiers 1..539 /organism="Homo sapiens" /db_xref="taxon:9606" /cclone_1lb="CI0154" /dev_stage="Adult" /note="Organ: colon.ins; Vector: puc18; site.1: Smal; site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters Patent application No. 196,716 - ludwig institute for cancer research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT	131 a 140 c 171 g 97 t
ORIGIN	
Query Match	32.3%; Score 32.6; DB 10; Length 539;
Best Local Similarity	56.9%; Pred.No. 2.8e+02;
Matches	56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
OY 7	cccccgtaggaagtgctggaatccttcaggaccacagaaggcggggcgagagg 66
Db 357	ccctcgacgaggagattgggtttgcgtcttcgaagtgcggctgccacgacgaacttggtc 298
OY 67	agtcctcctgagacctggttcctcctaagaacacag 101
Db 297	ATAGCCCATGCACGCACTGCGACTTGGCAGCCACAG 263
RESULT 9	
AL585199	
LOCUS	AL585199 685 bp mRNA linear EST 28-FEB-2001
DEFINITION	AL585199 BP Chicken Embryo Library Gallus gallus cdna clone
ACCSSION	R0502DD01, mRNA sequence.
VERSION	AL585199
KEYWORDS	AL585199.1 GI:13163932
SOURCE	EEST.
ORGANISM	Chicken.
	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
REFERENCE	Phasianidae; Gallus
AUTHORS	1 (bases 1 to 685)
TITLE	Murray,F. BP Chicken Embryo Library

JOURNAL COMMENT	Unpublished (2001)
	Contact: Frazer Murray
	Dept. Genomics and Bioinformatics
	Roslin Institute
	Roslin, Midlothian, EH25 9PS, UK
	Tel: +44 (0)131 527 4200
	Fax: +44 (0)131 440 0434
	Email: frazer.murray@bbsrc.ac.uk
	Seq primer: T7
FEATURES	Location/Qualifiers
source	1..685
	/organism="Gallus gallus"
	/db_xref="taxon:9031"
	/clone="ROS020D01"
	/clone_lib="BP_Chicken Embryo Library"
	/tissue_type="Embryo"
	/dev_stage="5 days old"
	/lab_host="DH10B"
	/note="vector: pBLUESCRIPT SK; Site_1: NotI; site_2: SalI;
	Cloned unidirectionally. Primer: Oligo dT. 5' adaptor
	sequence: 5' TCGAACCTCGAG 3'; 3' adaptor sequence: 5'
	GCGGCCGCCTTTTCTTTTTTTTTTTTTTTT 3'"
BASE COUNT	134 a 222 c 224 g 101 t 4 others
ORIGIN	
Query Match	31.9%; Score 32.2; DB 9; Length 685;
Best Local Similarity	65.7%; Pred. No.3.6e+02;
Matches	46; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Oy	17 ggagtcggacctggaactataggaccacaaggcgcgagggaggtcctctg 76
Dd	364 GGAGGAGGAGGGCCGACATGTGTGCGCCGACGAGGAGGCGGCGCTAACCCGCCG 423
Oy	77 gagcctgcgtg 86
Dd	424 GAGCTTGTTG 433
RESULT 10	
BG309093	735 bp mRNA linear EST 22-OCT-2001
LOCUS	HVSMEC0002B16f Hordeum vulgare seedling shoot EST library
DEFINITION	HVCDDNA0003 (/etiolated and unstressed) Hordeum vulgare cDNA clone
ACCESSION	BG309093
VERSION	BG309093.2 GI:16313793
KEYWORDS	EST.
SOURCE	barley.
ORGANISM	Hordeum vulgare
	Eukaryota; Viridiplantae; Hordeum vulgare; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae 1 (bases 1 to 735) Triticeae; Hordeum. Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton R.D., Gates,R. and Main,D. Development of a genetically and physically anchored EST resource for barley genomics: Morex unstressed seedling shoot cDNA library Unpublished (2001) On Feb 22, 2001 this sequence version replaced gi:13109940. Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Total hg bases = 300 Seq primer: AATTAAACCCTCACTAAGCG High quality sequence stop: 718. Location/Qualifiers 1..735
FEATURES	
source	1..735

/organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEC0002B16f"
 /clone_1lb="Hordeum vulgare seedling shoot EST library
 HVCNA0003 (Etiolated and unstressed)."
 /tissue_type="Seedling shoot"
 /lab_host="TJC121"

/note="Vector: lambdaZAP: Site 1: EcoRI; Site 2: XhoI;
 Seeds were surface sterilized then germinated under aseptic
 conditions in the dark at room temperature on filter paper
 with water, nystatin and ceftaxime in covered
 crystallization dishes. Five-day old seedling shoots were
 then harvested, total RNA was prepared, poly(A) RNA was
 purified, one primary unamplified cDNA library was made,
 and 1 million pfu were in vivo excised to give plasmid
 SK(-) cDNA phagemids. These steps were performed in the JY
 Close laboratory at the University of California,
 Riverside (Choi, Close, Fenton). Phagemids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Wang). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wang, Yu, Frisch, Henry, Simmons, Oates
 , Rambo, Main). The sequence has been trimmed to remove
 vector sequence and contains a minimum of 100 bases of
 phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wang R, Kiehlhofs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)."

BASE COUNT 189 a 156 c 264 g 116 t 10 others
 ORIGIN

Query Match 31.7%: Score 32; DB 10; Length 735;
 Best Local Similarity 62.5%: Pred. No. 4e+02;
 Matches 50; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 10 ctgtgagagatggctgtggaatctatgagcaccagaaggcgaggcgaggagtc 69
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 559 CTGCTGGCAGCGGCTGGCGGCGAGCGGAGCCACCGCGCTGAGCGGAGCGAGT 618
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 70 cctctggagccctgtgccc 89
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 619 TTGGCTGGCTGTGTGGCCCC 638

RESULT 11 382 bp mRNA linear EST 18-AUG-2000
 BE600720/c
 LOCUS P11_91_G10_b1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
 DEFINITION
 mRNA sequence.
 BE600720
 VERSION BE600720.1 GI:9855905
 KEYWORDS EST.
 SOURCE sorghum.
 ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 382)
 AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
 ,L.H.

TITLE An EST database from Sorghum: pathogen-induced plants
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860

Fax: 706 542 1805
 Email: mmp@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 382
 POLYA-No.

FEATURES
 Source location/Qualifiers

1..382
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_1lb="Pathogen induced 1 (P11)"
 /note="Organ: Anthracnose-infected leaves from
 two-week-old sorghum plants 48 hr after inoculation;
 Vector: plasmid II from lambda Zap II; Site_1: XhoI;
 Site_2: EcoRI; Two-week-old sorghum plants (Brix 623
 cultivar) were infected with pathogen (isolate PRM421 of
 Colletotrichum graminicola, which is a sorghum isolate).
 RNA was prepared from infected leaves harvested from 45
 seedlings 48 hours after inoculation. Note: young
 seedlings (2 weeks old) exhibit juvenile resistant
 reaction, which is an incompatible interaction. As they
 grow older (4 weeks or older), plants resume susceptibility
 to anthracnose disease. The library was made from poly-A
 RNA in the cloning vector lambda Zap II. Clones to be
 sequenced were prepared by mass excision. WARNING: While
 most or all ESTs are expected to derive from the host
 plant, no effort was made to eliminate ESTs deriving from
 the pathogen.."

BASE COUNT 67 a 134 c 90 g 91 t
 ORIGIN

Query Match 31.5%: Score 31.8; DB 10; Length 382;
 Best Local Similarity 59.3%: Pred. No. 4.2e+02;
 Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 11 ttgtgagagatggctgtggaatctatgagcaccagaaggcgaggcgaggagtc 70
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 103 TGATGATGTGTGGGCTGGCTGCTATGCGGAGCGGAGCGAGGAGCGAGAG 44
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 71 ctcttgagccctgtgctcctagaagccacg 101
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 43 GGAGAAATGCTGTGTGCCGATGCCCATG 13

RESULT 12 463 bp mRNA linear EST 18-AUG-2000
 BE600139/c
 LOCUS P11_79_G07_b1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
 DEFINITION
 mRNA sequence.
 BE600139
 VERSION BE600139.1 GI:9855212
 KEYWORDS EST.
 SOURCE sorghum.
 ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 463)
 AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
 ,L.H.

TITLE An EST database from Sorghum: pathogen-induced plants
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Department of Botany
 The University of Georgia
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 542 1805
 Email: mmp@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions

Query Match	27.38;	Score 27.6;	DB 4;	Length 6822;
Best Local Similarity	56.78;	Pred. No. 14;		

Sequence 1, Application US/083244A3B
Patent No. 5654170

GENERAL INFORMATION:

APPLICANT: KLINGER, KATHERINE W.
APPLICANT: LANDES, GREGORY M.
APPLICANT: BURN, TIMOTHY C.
APPLICANT: CONNORS, TIMOTHY D.
APPLICANT: DACKOWSKI, WILLIAM R.
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, PENG

TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/323.4438
: FILING DATE: 12-OCT-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ludwig, S. Peter
: REGISTRATION NUMBER: 25,351
: REFERENCE/DOCKET NUMBER: 0372/0A462
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 527-7700
: TELEFAX: (212) 753-6237
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31571 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: CLONE: PDI GENOMIC
: US-08-323-4438-1

Query Match
Best Local Similarity 58.8%; Score 27.2; DB 1; Length 31571;
Matches 47; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 5 ccccccgtgtagaggtggtggaatctatgagccaccagagggcgagggcgaggg 64
Db 16428 CCTGCTGGGTACTGCTGGGAGGCTGACACCTGGGAGTAGAGGCCCGTGGCAGGAGG 16487
QY 65 ggaagtcctcctcgagcctgg 84
Db 16488 TGAGGCTCGGGCTGCTGG 16507

RESULT 6
US-08-658-136-2
: Sequence 2, Application US/08658136
: Patent No. 6071717
: GENERAL INFORMATION:
: APPLICANT: KLINGER, KATHERINE W
: APPLICANT: LANDES, GREGORY M
: APPLICANT: BURN, TIMOTHY C
: APPLICANT: CONNORS, TIMOTHY D
: APPLICANT: DACKOWSKI, WILLIAM
: APPLICANT: GERMINO, GREGORY
: APPLICANT: QIAN, FENG
: TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENZYME CORPORATION
: STREET: ONE MOUNTAIN ROAD
: CITY: FRAMINGHAM
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 01701
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/658,136
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: LASSEN, ELIZABETH
: REGISTRATION NUMBER: 31,845

US-08-658-136-1
```

```

: REFERENCE/DOCKET NUMBER: GEN4-17.8
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 508-872-8400
: TELEFAX: 508-872-5415
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 53526 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-658-136-2

Query Match
Best Local Similarity 58.8%; Score 27.2; DB 3; Length 53526;
Matches 47; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 5 ccccccgtgtagaggtggtggaatctatgagccaccagagggcgagggcgaggg 64
Db 17414 CCTGCTGGGTACTGCTGGGAGGCTGACACCTGGGAGTAGAGGCCCGTGGCAGGAGG 17473
QY 65 ggaagtcctcctcgagcctgg 84
Db 17474 TGAGGCTCGGGCTGCTGG 17493

RESULT 7
US-08-658-136-1
: Sequence 1, Application US/08658136
: Patent No. 6071717
: GENERAL INFORMATION:
: APPLICANT: KLINGER, KATHERINE W
: APPLICANT: LANDES, GREGORY M
: APPLICANT: BURN, TIMOTHY C
: APPLICANT: CONNORS, TIMOTHY D
: APPLICANT: DACKOWSKI, WILLIAM
: APPLICANT: GERMINO, GREGORY
: APPLICANT: QIAN, FENG
: TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENZYME CORPORATION
: STREET: ONE MOUNTAIN ROAD
: CITY: FRAMINGHAM
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 01701
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/658,136
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: LASSEN, ELIZABETH
: REGISTRATION NUMBER: 31,845
: REFERENCE/DOCKET NUMBER: GEN4-17.8
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 508-872-8400
: TELEFAX: 508-872-5415
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 53577 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-658-136-1
```

Query Match 26.9%; Score 27.2; DB 3; Length 53577;
Best Local Similarity 58.8%; Pred. No. 23;
Matches 47; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 5 cccccctggtgagagctgggaattctatggcaccacaagggcgaggcggaag 64
DB 17413 CTTCTGCGTACTGCTGGGAGGCTGACACTGGGAGAGTACAGCCCGTGGCAGGAG 17472

OY 65 ggaagtcctctggagcctg 84
DB 17473 TGAGCCCTCGGCTCTCGG 17492

RESULT 8
US-08-451-947-9/c
Sequence 9, Application US/08451947
Patent No. 5702906
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUTROTROPIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451.947
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 666P2C102
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: Linear
US-08-451-947-9

Query Match 26.5%; Score 26.8; DB 1; Length 1190;
Best Local Similarity 57.0%; Pred. No. 20;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 11 tggtagagagtggtggaattctatggcaccacaagggcgaggcggaagtc 70
US-08-928-694-9/c

DB 469 TGTGAGAGATGGGCTTAGACATCCATGACACACTGGGAGAGGAAATGAGGGGATG 410
OY 71 ctcttgagacctgggacctagaagc 96
DB 409 CGAGGAGGCTGGGGAGCAGGAGC 384

RESULT 9
US-08-424-826A-9/c
Sequence 9, Application US/08424826A
Patent No. 5830858
GENERAL INFORMATION:
APPLICANT: Rosenthal, Arnon
TITLE OF INVENTION: NOVEL NEUTROTROPIC FACTOR
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424.826A
FILING DATE: 19-APR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240387
FILING DATE: 10-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 25-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Phd., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0666P1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-424-826A-9

Query Match 26.5%; Score 26.8; DB 2; Length 1190;
Best Local Similarity 57.0%; Pred. No. 20;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 11 tggtagagagtggtggaattctatggcaccacaagggcgaggcggaagtc 70
DB 469 TGTGAGAGATGGGCTTAGACATCCATGACACACTGGGAGGAAATGAGGGGATG 410
OY 71 ctcttgagacctgggacctagaagc 96
DB 409 CGAGGAGGCTGGGGAGCAGGAGC 384

RESULT 10
US-08-928-694-9/c

```
Sequence 9, Application US/08928694
Patent No. 6037320
GENERAL INFORMATION:
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Minipain (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,694
FILING DATE: 12-Sep-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/451947
FILING DATE: 26-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013
FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Phd., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0666P2C1D2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/952-9881
TELEFAX: 650/952-8674
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-928-694-9

Query Match 26.5%; Score 26.8; DB 3; Length 1190;
Best Local Similarity 57.0%; Pred. No. 20;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Oy 11 tggtagagagctgggctgggaattctatgggaccccaagagggcggggagggagtc 70
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 469 TGTGAGAGATGGGTTGAGACTCCAAATGACACACTGGGAGAGAGAAATGAGGGGATG 410
Oy 71 ctctgagagctggtgcctagaagc 96
    |  |||||||  |  |||||
Db 409 CGAGAGGAGCCTGGGAGGAGCAGCAGC 384

RESULT 11
PCT-US91-06950-9/c
Sequence 9, Application PC/TUS9106950
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
```

```
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: pain (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06950
FILING DATE: 19910924
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 666P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US91-06950-9

Query Match 26.5%; Score 26.8; DB 5; Length 1190;
Best Local Similarity 57.0%; Pred. No. 20;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Oy 11 tggtagagagctgggctgggaattctatgggaccccaagagggcggggagggagtc 70
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 469 TGTGAGAGATGGGTTGAGACTCCAAATGACACACTGGGAGAGAGAAATGAGGGGATG 410
Oy 71 ctctgagagctggtgcctagaagc 96
    |  |||||||  |  |||||
Db 409 CGAGAGGAGCCTGGGAGGAGCAGCAGC 384

RESULT 12
US-09-211-417-2/c
Sequence 2, Application US/09211417A
Patent No. 6177254
GENERAL INFORMATION:
APPLICANT: Ratner, Jerome B
APPLICANT: Whitehead, Clark M
TITLE OF INVENTION: NUCLEOLUS AUTOANTIGENIC MARKER FOR SYSTEMIC LUPUS
TITLE OF INVENTION: ERHHEMATOSUS
TITLE OF INVENTION: Genbank
FILE REFERENCE: UCC1
CURRENT APPLICATION NUMBER: US/09/211,417A
CURRENT FILING DATE: 1998-12-15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 3286
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Nucleic Acid Sequence of ASE-1
```



```

: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/290,665A
: FILING DATE: 15-AUG-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: RICHARD W. BORK
: REGISTRATION NUMBER: 36,459
: REFERENCE/DOCKET NUMBER: 2026-4116
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 758-4800
: TELEFAX: (212) 751-6849
: TELEX: 421792
: INFORMATION FOR SEQ ID NO: 140:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 573 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: ORIGINAL SOURCE:
: ORGANISM: homosapiens
: INDIVIDUAL ISOLATE: 28
: US-08-290-665A-140

```

```

Query Match      26.3%; Score 26.6; DB 2; Length 573;
Best Local Similarity 56.2%; Pred. No. 21;
Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 5 cccccctgtaggagctggctggaalctatgggcacccagaggcgcgagcgaag 64
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 359 CCCAATTGCGGACCTCCGCGGGATCATTTGGGCCCCAAGACGTCGAGAGCGCGCG 300
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 65 ggaatcctcctgagagcctgtagccctaga 93
   111 111 111 111 111 111 111 111 111 111 111 111 111
Db 299 GCGGACAGAGAGCCCAACTGCCCCACCCACA 271
   111 111 111 111 111 111 111 111 111 111 111 111 111

```

Search completed: October 3, 2002, 16:21:45
 Job time: 16945 sec

PN W0200177305-A2.
XX
XX PD 18-OCT-2001.
XX
XX PF 06-APR-2001; 2001WO-SE00765.
XX
XX PR 07-APR-2000; 2000US-195665P.
XX
XX PA (AREX-) AREXIS AB.
XX
XX PI Andersson L, Luthman H, Marklund S;
XX WPI: 2001-657170/75.
DR P-PSDB; OQB47679.
DR

PT New variants of human AMP-activated protein kinase gamma3 subunit
PT associated with a metabolic disease e.g. diabetes or obesity and method
PT for determining a risk estimate of diseases in subject by detecting the
PT variant -
XX
PS Disclosure: Fig 5; 25pp; English

CC This sequence represents the full length cDNA encoding the human
CC AMP-activated protein kinase gamma 3 subunit (PRKG3). Detecting
CC the presence of the PRKG3 DNA, or a variant, is useful in determining
CC a risk estimate of a metabolic disease, such as diabetes or obesity,
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon
CC 3 variation may be a substitution of a G for a C at nucleotide 320,
CC resulting in the amino acid substitution P71A; in exon 4 variation may
CC be a substitution of a T for a C at nucleotide 550; and in exon 10
CC variation may be a substitution of a T for a C at nucleotide 1037,
CC resulting in the amino acid substitution R340W. There may also be
CC nucleotide variation in intron 6. The numbering of these
CC variations is based on the full length cDNA as given, rather than on
CC position 1 of the open reading frame.

S0 Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;

Query Match	100.0%	Score 1647	DB 22	Length 1647
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1647	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

[illegible]

Dh	421	gctvgccacagvgahtlccccagccacagavgcctvgvgagtvgagctlaagavgcctgtctga	480
Qy	481	agagavgcctvgccctgtgtccctgtlcccccagagcccccattcccacagctvgvgatga	540
Dh	481	agagavgcctvgccctgtgtccctgtlcccccagagcccccattcccacagctvgvgatga	540
Qy	541	cgaactvgvgaaaacccvggccccagatctacatgcgtctcaatgcagvgaacacatgtcta	600
Dh	541	cgaactvgvgaaaacccvggccccagatctacatgcgtctcaatgcagvgaacacatgtcta	600
Qy	601	cgaatgcataatgcacataagcttccaaagctagcatctctgacacccatgtcttgagatctcaaaa	660
Dh	601	cgaatgcataatgcacataagcttccaaagctagcatctctgacacccatgtcttgagatctcaaaa	660
Qy	661	gacctctctgtctctvgctvgccacagvtgtcvgggcagccccctcatatvggacagcaaaa	720
Dh	661	gacctctctgtctctvgctvgccacagvtgtcvgggcagccccctcatatvggacagcaaaa	720
Qy	721	gcagagctcttvgvgvgatgtctacacatacagacttcaatccctgtgtctgatactga	780
Dh	721	gcagagctcttvgvgvgatgtctacacatacagacttcaatccctgtgtctgatactga	780
Qy	781	caggtccccctvgctccagatctatagaaattgaaacaataagaattgagacctvgagga	840
Dh	781	caggtccccctvgctccagatctatagaaattgaaacaataagaattgagacctvgagga	840
Qy	841	gactacccctvgcaagvgctgcttcaaaagccctgtgtctcatatctccctaaatgataagctgt	900
Dh	841	gactacccctvgcaagvgctgcttcaaaagccctgtgtctcatatctccctaaatgataagctgt	900
Qy	901	tgaagctgtctacacccctcatcaagaaacggaatccatcgcctgtgcctgtcttgaacccggt	960
Dh	901	tgaagctgtgtctacacccctcatcaagaaacggaatccatcgcctgtgcctgtcttgaacccggt	960
Qy	961	gtccagagcaagctgactccacatctccacacacaaacgaagctgtctcaagtctctgacatctt	1020
Dh	961	gtccagagcaagctgactccacatctccacacacaaacgaagctgtctcaagtctctgacatctt	1020
Qy	1021	tggttccctgtcgtcccgccgcccctctctctcaacgacatactcaagatttvgcatcgg	1080
Dh	1021	tggttccctgtcgtcccgccgcccctctctctcaacgacatactcaagatttvgcatcgg	1080
Qy	1081	gacatctccgagacttgctgtgtgtgtcgtcvgagacagacccaatccatgtcaactvgacat	1140
Dh	1081	gacatctccgagacttgctgtgtgtgtgtcgtcvgagacagacccaatccatgtcaactvgacat	1140
Qy	1141	cttttvgaacccgvgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1200
Dh	1141	cttttvgaacccgvgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1200
Qy	1201	ccctcatctccggtctgaatgatactacatcctgtcgtccacagcaaaacttaaacacacttga	1260
Dh	1201	ccctcatctccggtctgaatgatactacatcctgtcgtccacagcaaaacttaaacacacttga	1260
Qy	1261	catvgagtvgvgaaagacccctvgagagacagaaactaagatctcvgvgvgaggtctcttctcgt	1320
Dh	1261	catvgagtvgvgaaagacccctvgagagacagaaactaagatctcvgvgvgaggtctcttctcgt	1320
Qy	1321	ccagccccacagagacttvgvgggaagtgaatcagacagatgtctcgtvgvgccaggtctaacag	1380
Dh	1321	ccagccccacagagacttvgvgggaagtgaatcagacagatgtctcgtvgvgccaggtctaacag	1380
Qy	1381	gcttvgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1440
Dh	1381	gcttvgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1440
Qy	1441	tcagagcacctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1500
Dh	1441	tcagagcacctgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1500
Qy	1501	tctctcaatcccaagcacctgtcacaacttgaagccaatgaagvgggaactvggaactcagc	1560
Dh	1501	tctctcaatcccaagcacctgtcacaacttgaagccaatgaagvgggaactvggaactcagc	1560


```

OY 1561 ctccatcttccccaccacalltgcgtgctcagctatgattcaggtaggtctctgccc 1620
DB 1561 ctccatcttccccaccacacccattgctggtcagctatgattcaggtaggtctgccc 1620
OY 1621 ggcacatgacacccagctcttaagcttc 1647
DB 1621 ggcacatgacacccagctcttaagcttc 1647

RESULT 2
AADD3320
ID AADD3320 standard; cDNA: 2115 BP.
XX
AC AADD3320;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key 1..1395 Location/Qualifiers
FT CDS /tag= a
FT /product= "Human complete Prkag3 protein"
FT
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000: 2000WO-EP09896.
XX
PR 10-SEP-1999: 99EP-0402236.
PR 18-MAY-2000: 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Tamnucellil N, Gellin J, Le Roy P, Chardon P;
XX
DR MPI: 2001-244810/25.
DR P-PSDB: AAE00223.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 12: Page 65-68; 71pp: English.
XX
XX The present sequence is a cDNA encoding human adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC useful as therapeutic for treating carbohydrate metabolism disorders such
CC as diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.

```

```

CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;

Query Match      88.2%; Score 1453; DB 22; Length 2115;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 35; Indels 7; Gaps 2;

OY 95 atgagcttccttagagcaaaaacagcaatcatgacctacacagctgtgacagcagc 154
DB 1 atgagcttccttagagcaaaaacagcaatcatgacctacacagctgtgacagcagc 60
OY 155 tcagaaagaatccgttgggaacgagggccaaagccttgaatgagacaaagcagaagtcg 214
DB 61 tcagaaagaatccgttgggaacgagggccaaagccttgaatgagacaaagcagaagtcg 120
OY 215 gtggaggaaggaggccaccaggtcagggggaaggtcccggtccagggcagctgtgag 274
DB 121 gtggaggaaggaggccaccaggtcagggggaaggtcccggtccagggcagctgtgag 180
OY 275 tcacccgggtctgagggccacatctccccaagacacaccccttgctcctaaagtatctgtcc 334
DB 181 tcacccgggtctgagggccacatctccccaagacacaccccttgctcctaaagtatctgtcc 240
OY 335 ggggttgggcactccaccaaagggttgggactgtccctccctctgactgtacagctcagct 394
DB 241 ggggttgggcactccaccaaagggttgggactgtccctccctctgactgtacagctcagct 300
OY 395 gcaagctccagacagatgatgtgagctgtgccaaggagttccaggaacagaagccttgg 454
DB 301 gcaagctccagacagatgatgtgagctgtgccaaggagttccaggaacagaagccttgg 360
OY 455 gagtgtgaagtagaaggcctgtctggaagaagagcctgtgctgtgtcccgagagcc 514
DB 361 gagtgtgaagtagaaggcctgtctggaagaagagcctgtgctgtgtgtcccgagagcc 420
OY 515 ccatttcccaagctggcgtgggatacgaactgtgggaaacccggcgccagatctcaatg 574
DB 421 ccatttcccaagctggcgtgggatacgaactgtgggaaacccggcgccagatctcaatg 480
OY 575 cgcctcatgaggaacacacccgtctacgattgccaatgcaactagctccaaagtatcatc 634
DB 481 cgcctcatgaggaacacacccgtctacgattgccaatgcaactagctccaaagtatcatc 540
OY 635 ttcgaaccatgtctgagatcaagaaggcctcttctgtctgtgtgccaagtggtgcgg 694
DB 541 ttcgaaccatgtctgagatcaagaaggcctcttctgtctgtgtgccaagtggtgcgg 600
OY 695 gaagccctctatggaagcaagaagaagacgttgttggggatgtcgaaccatcaatgac 754
DB 601 gaagccctctatggaagcaagaagaagacgttgttggggatgtcgaaccatcaatgac 660
OY 755 ttcacatctgtctgcatcgctacacaggttccccccgtgtccaagatctatagattgaa 814
DB 661 ttcacatctgtctgcatcgctacacaggttccccccgtgtccaagatctatagattgaa 720
OY 815 caacataagattgagacactggaaggagatctacactgaaggtgtcttaagcctctgtc 874
DB 721 caacataagattgagacactggaaggagatctacactgaaggtgtcttaagcctctgtc 780
OY 875 tccacatctccaaagatagactgtttaaagctgtctacacccctatcaagaacgggtc 934
DB 781 tccacatctccaaagatagactgtttaaagctgtctacacccctatcaagaacgggtc 840
OY 935 catgcccgtctgtctctgacccgggtgtcaggaacgtaactccaatctcaacacaaa 994
DB 841 catgcccgtctgtctctgacccgggtgtcaggaacgtaactccaatctcaacacaaa 900

```

OY	995	cgccctgcctcaagatctcctgcacacacttggttctccctgcctgcgcccgggcctctctctctac	1054
Db	901	cgccctgcctcaagatctcctgcacacacttggttctccctgcctgcgcccgggcctctctctctac	960
OY	1055	cgacactatccaaagatttgaggcatcggccacatcccgagacttggctggtgctctgagaca	1114
Db	961	cgacactatccaaagatttgaggcatcggccacatcccgagacttggctggtgctctgagaca	1020
OY	1115	gacacacatccctgagctgcacacgagacacactcttgagacccgctgctgtctcgcactgctctg	1174
Db	1021	gacacacatccctgagctgcacacgagacacactcttgagacccgctgctgtctcgcactgctctg	1080
OY	1175	gtcaacagaaatgctgtcagcgtgcgtgggcctctatctccgccttggatgattcaactgct	1234
Db	1081	gtcaacagaaatgctgtcagcgtgcgtgggcctctatctccgccttggatgattcaactgct	1140
OY	1235	gcccagagaaacactcaacacacacactcgcagacagatgtgagaaagcccttaggcagagagaca	1294
Db	1141	gcccagagaaacactcaacacacacacactcgcagacagatgtgagaaagcccttaggcagagagaca	1200
OY	1295	ctatgtcctcgtgagagagatcctcttccctgcgcagaccccacgaaagcttgagggaaagtcatgcac	1354
Db	1201	ctatgtcctcgtgagagagatcctcttccctgcgcagaccccacgaaagcttgagggaaagtcatgcac	1260
OY	1355	agagcttgctcctggagagcagcgtgacacaaagcctggtgtctagatgagaaagaaaccaaagatctctg	1414
Db	1261	agagatgtctcctggagagcagcgtgacacaaagcctggtgtctagatgagaaagaaaccaaagatctctg	1320
OY	1415	ggcgtggctctcccccctccgcagacatctctccagcactgctgtgctcagccctgcctgcacgat	1474
Db	1321	ggcgtggctctcccccctccgcagacatctctccagcactgctgtgctcagccctgcctgcacgat	1380
OY	1475	ggccctcgggggcctgagaaagatctcgatctcctcaatcccaagccaactgcacacctgagaagc	1534
Db	1381	ggccctcgggggcctgagaaagatctcgatctcctcaatcccaagccaactgcacacctgagaagc	1439
OY	1535	caatgaaagagaaatcgagaaatcagctcatcattcttcccacccaacccaattgtgtgtcag	1594
Db	1440	caatgaaagagaaatcgagaaatcagctcatcattcttcccacccaacccaattgtgtgtcag	1493
OY	1595	ctatgtatctcaagttagcctctgcctcctgcgagcactgacacagcact	1637
Db	1494	ctatgtatctcaagttagcctctgcagcctcccaaatatgaccttgcctg	1536

RESULT	3	
AD003296		
ID	AD003296	standard; DNA; 2109 BP.
XX		
AC	AD003296;	
XX		
DT	13-JUN-2001	(first entry)
XX		
DE	Human AMPK gamma subunit muscle-specific isoform, PRKAG3 CDNA.	
XX		
RM	Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;	
KM	PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;	
XX	genetic testing; carbohydrate metabolism disorder; skeletal muscle;	
KW	cystathione beta synthase; CBS; cardiact; gene therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	5'UTR	1..471
FT		/*tag= a
FT	CDS	472..1389
FT		/*tag= b
FT		/product= "Human Prkag3 protein"
FT	3'UTR	1390..2109
FT		/*tag= c
XX		
PN	MO200120003-A2.	

XX	22-MAR-2001.
PD	
XX	
PF	11-SEP-2000; 2000MO-EP09896.
XX	
PR	10-SEP-1999; 99EP-0402236.
XX	
PR	18-MAY-2000; 2000EP-0401388.
XX	
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.
XX	
PA	(ANDE/) ANDERSSON L.
XX	
PA	(LOOF/) LOOFT C.
XX	
PI	(KALM/) KALM E.
XX	
PI	Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Galliard C,
XX	
PI	Iannuccielli N, Gellin J, Le Roy P, Chardon P,
XX	
DR	WPI: 2001-244810/25.
XX	
DR	P-PSDB: AAE00221.
XX	
PT	New variants of the gamma subunit of vertebrate adenosine
XX	
PT	monophosphate-activated kinase for diagnosis or treatment of disorders
XX	
PT	associated with energy metabolism such as diabetes, obesity, and
XX	
PT	myopathy
XX	
PS	
XX	
CC	Claim 12; Fig 2; 71pp: English.
XX	
CC	The present sequence is a CDNA encoding human adenosine monophosphate
XX	
CC	(AMP) activated kinase (AMPK) gamma subunit muscle-specific isoform,
XX	
CC	PRKAG3. Mutation in Prkag3 results in an altered regulation of
XX	
CC	carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
XX	
CC	useful as therapeutic for treating carbohydrate metabolism disorders such
XX	
CC	as diabetes, obesity, and disorders associated with muscle metabolism
XX	
CC	such as myopathy and cardiovascular diseases, to modulate AMPK
XX	
CC	activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX	
CC	and its functionally altered mutants are useful for the diagnostic
XX	
CC	evaluation, genetic testing and prognosis of a metabolic disorder,
XX	
CC	preferably a carbohydrate metabolism disorder. Primers that can detect
XX	
CC	a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX	
CC	useful for detecting a dysfunction of carbohydrate metabolism resulting
XX	
CC	from the expression of a functionally altered allele of PRKAG3.
XX	
CC	Transgenic animal and host cell transformed with PRKAG3 or a
XX	
CC	heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX	
CC	screening compounds able to modulate AMPK activity. Nucleic acid
XX	
CC	encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
XX	
CC	in a sequence encoding the first cystathione beta synthase (CBS) domain
XX	
CC	of PRKAG3 and is useful in gene therapy.
XX	
CC	
XX	
SQ	Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other:
XX	
Query Match	87.9%; Score 1447; DB 22; Length 2109;
Best Local Similarity	97.3%; Pred. NO. 0;
Matches 1495; Conservative	0; Mismatches 35; Indels 7; Gaps 2;
101	ttcctagagcaagaacaacagcagctctcatggtccatcaccagctgttaccagcagctccagaa 160
DB	1 ttcctagagcaagaacaacagcagctctcatggtccatcaccagctgttaccagcagctccagaa 60
161	agaatccgtgtggaacacgagggccaaagccttgagatggaacaaggcagaagtcggttgag 220
DB	61 agaatccgtgtggaacacgagggccaaagccttgagatggaacaaggcagaagtcggttgag 120
221	gaagggagagccacacgctcagtgagggaaggtccccggtcccaaggccagctgtgagtcacg 280
DB	221 gaagggagagccacacgctcagtgagggaaggtccccggtcccaaggccagctgtgagtcacg 280
281	ggagctggagagccacacgtccccaaggccacaccccttggtctcaagctgagctccggcggtg 340
DB	181 ggagctggagagccacacgtccccaaggccacaccccttggtctcaagctgagctccggcggtg 240
341	ggactccaccaacaacgggttggaatgtcctccctcttgactgtacagctcagctgtcaagc 400
DB	241 ggactccaccaacaacgggttggaatgtcctccctcttgactgtacagctcagctgtcaagc 300

QY 401 tccagacagatgagtgagagctgcccagagcttccagcacagagagcctggagtg 460
 |||||
 Db 301 tccagacagatgagtgagagctgcccagagcttccagcacagagagcctggagtg 360
 QY 461 gagctgaagagcctgctggaagagagagcctgcccgtgctgttccccgagagcccat 520
 |||||
 Db 361 gagctgaagagcctgctggaagagagagcctgcccgtgctgttccccgagagcccat 420
 QY 521 cccaaagctggagtgagtgacgaactgcccgaaccggagccagatcaactgagcttc 580
 |||||
 Db 421 cccaaagctggagtgagtgacgaactgcccgaaccggagccagatcaactgagcttc 480
 QY 581 atgcagagagacacacctgtgacgaatgccaatggaactagctccaagctagtcctgcag 640
 |||||
 Db 481 atgcagagagacacacctgtgacgaatgccaatggaactagctccaagctagtcctgcag 540
 QY 641 accatgctggagatcaagaagagcctcttctgctcgtggagccaaagagtgctggagcagcc 700
 |||||
 Db 541 accatgctggagatcaagaagagcctcttctgctcgtggagccaaagagtgctggagcagcc 600
 QY 701 cctctatggagaaagaagaagcagagcttctgggagtgctgacacacacacacacacacac 760
 |||||
 Db 601 cctctatggagaaagaagaagcagagcttctgggagtgctgacacacacacacacacacac 660
 QY 761 ctggtgctgacatgctactacagagctccccctggtccagatcagatgagatggaacaacat 820
 |||||
 Db 661 ctggtgctgacatgctactacagagctccccctggtccagatcagatgagatggaacaacat 720
 QY 821 aagatgagacctggagagagatcctacatgaagctgcttcaagcctctgcttccatc 880
 |||||
 Db 721 aagatgagacctggagagagatcctacatgaagctgcttcaagcctctgcttccatc 780
 QY 881 tctccaaatgagtgagctgttggagagctgtctacacccctcacaagaacagagatccatcgc 940
 |||||
 Db 781 tctccaaatgagtgagctgttggagagctgtctacacccctcacaagaacagagatccatcgc 840
 QY 941 ctgacctgttctgacccggtgtcagagcaagtaclccacacccctcacaagaacagcctg 1000
 |||||
 Db 841 ctgacctgttctgacccggtgtcagagcaagtaclccacacccctcacaagaacagcctg 900
 QY 1001 ctcaagttccctgacacatcttgggttccctgctgcccggagcccttcccttaccagacat 1060
 |||||
 Db 901 ctcaagttccctgacacatcttgggttccctgctgcccggagcccttcccttaccagacat 960
 QY 1061 atccaagaattgagcactgacacatctccgaagacttgctggtgagctggagagagagagcacc 1120
 |||||
 Db 961 atccaagaattgagcactgacacatctccgaagacttgctggtgagctggagagagagcacc 1020
 QY 1121 atcctgacatgacatgacatcttctggaaccggagctgtgtctgacatgacatgacatgac 1180
 |||||
 Db 1021 atcctgacatgacatgacatcttctggaaccggagctgtgtgtctgacatgacatgacatgac 1080
 QY 1181 gaatgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1240
 |||||
 Db 1081 gaatgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1140
 QY 1241 caaaactacaac 1300
 |||||
 Db 1141 caaaactacaac 1200
 QY 1301 ctggagagagagtgcttcttctgcccagccccagagagcttggggagagtgatcgacagatc 1360
 |||||
 Db 1201 ctggagagagagtgcttcttctgcccagccccagagagcttggggagagtgatcgacagatc 1260
 QY 1361 gctcggagagagtgacacagagctgt 1420
 |||||
 Db 1261 gctcggagagagtgacacagagctgt 1320
 QY 1421 gctcctctctcgacatcctctcagagcactgtgtcagagccctgtgcatcgatgctccctc 1480
 |||||
 Db 1321 gctcctctctcgacatcctctcagagcactgtgtcagagccctgtgcatcgatgctccctc 1380

QY 1481 ggggctgagaagatcgtatcctcattcccaagccacactgacacactggagagccatgga 1540
 |||||
 Db 1381 ggggcttggagaagatcgtatcctcattcccaagccacactgacac-cggagagccaatgga 1439
 QY 1541 aggaactggagaagactcagcccttcatcttccccaccccatctgtgtgtcagctatgga 1600
 |||||
 Db 1440 aggaactgaga-----acagcttcatcttccccaaccccaatttgcgtgtcagctatgga 1493
 QY 1601 ttcaagtagagctgtgctccttgggagccatgacacagcct 1637
 |||||
 Db 1494 ttcaagcttcttgcagcttcccaaatgtgcttgcct 1530

RESULT 4

AAD03321

ID AAD03321 standard; DNA; 2022 BP.

XX

AC

AAD03321;

DT

13-JUN-2001 (first entry)

XX

DE

Sus scrofa PRKAG3 splice variant DNA.

XX

KW

PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW

genetic testing; carbohydrate metabolism disorder; skeletal muscle;

KW

cystathione beta synthase; CBS; cardiact; gene therapy; ds.

XX

OS

Sus scrofa.

XX

FH

Key

FT

CDS

1..1545

/*tag= a

/product= "Sus scrofa Prkag3 splice variant"

WT0200120003-A2.

PD

22-MAR-2001.

XX

PF

11-SEP-2000; 2000MO-EP09896.

XX

PR

10-SEP-1999; 99EP-0402236.

PR

18-MAY-2000; 2000EP-0401388.

XX

PA

(INRG) INRA INST NAT RECH AGRONOMIQUE.

PA

(ANDE/) ANDERSSON L.

PA

(LOOF/) LOOF C.

PA

(KALM/) KALM E.

PI

Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI

Iannuccelli N, Gellin J, Le Roy P, Chardon P;

DR

WPI: 2001-244810/25.

DR

P-PSDB; AAE00224.

Claim 12; Page 69; 71pp; English.

The present sequence is pig adenosine monophosphate (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice variant DNA. Prkag3 gene is located in the RN locus of chromosome 15. Mutation in Prkag3 results in an altered regulation of carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is useful as therapeutic for treating carbohydrate metabolism disorders such as diabetes, obesity, and disorders associated with muscle metabolism such as myopathy and cardiovascular diseases, to modulate AMPK activity, and for restoring a normal AMPK function. PRKAG3 sequence and its functionally altered mutants are useful for the diagnostic evaluation, genetic testing and prognosis of a metabolic disorder,


```

FT CDS 1..1395
FT /*tag= a
FT /product= *Sus scrofa complete Prkag3 protein*
XX
XX MO200120003-A2.
XX
XX 22-MAR-2001.
XX
XX 11-SEP-2000: 2000WO-EP09896.
XX
XX 10-SEP-1999: 99EP-0402236.
XX
XX 18-MAY-2000: 2000EP-0401388.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFF C.
XX (KALM/) KALM E.
XX
XX Andersson L, Looff C, Kalm E, Milan D, Robic A, Rogel-Gallard C;
XX Iannucci L N, Gellin J, Le Roy P, Chardon P;
XX WPI: 2001-244810/25.
XX P-PSDB: AAE00222.
XX
XX New variants of the gamma subunit of vertebrate adenosine
XX monophosphate-activated kinase for diagnosis or treatment of disorders
XX associated with energy metabolism such as diabetes, obesity, and
XX myopathy
XX
XX Claim 12: Page 62-64: 71pp: English.
XX
XX The present sequence is a cDNA encoding pig adenosine monophosphate
XX (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
XX complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
XX 15. Mutation in Prkag3 results in an altered regulation of carbohydrate
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX therapeutic for treating carbohydrate metabolism disorders such as
XX diabetes, obesity, and disorders associated with muscle metabolism
XX such as myopathy and cardiovascular diseases, to modulate AMPK
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX and its functionally altered mutants are useful for the diagnostic
XX evaluation, genetic testing and prognosis of a metabolic disorder,
XX preferably a carbohydrate metabolism disorder. Primers that can detect
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX useful for detecting a dysfunction of carbohydrate metabolism resulting
XX from the expression of a functionally altered allele of PRKAG3.
XX Transgenic animal and host cell transformed with PRKAG3 or a
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX screening compounds able to modulate AMPK activity. Nucleic acid
XX encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
XX in a sequence encoding the first cystathione beta synthase (CBS) domain
XX of PRKAG3 and is useful in gene therapy.
XX
XX Sequence 1873 BP: 382 A: 580 C: 535 G: 376 T: 0 other:

```

```

Query Match 69.2%; Score 1140.4; DB 22; Length 1873;
Best Local Similarity 85.7%; Pred. No. 1,1e-287;
Matches 1315; Conservative 0; Mismatches 211; Indels 8; Gaps 4;

```

```

OY 95 atgagcttcttagaagaagaacagcagctcgtcattcaccagctgtgaccagcagc 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 atgagcttcttagaagaagaagcgtcattcaccagctcgtcattcaccagcagc 60
OY 155 tcaagaagaatcgttggaagcgaagcgaagccttgagatggaacaagcagaatcgc 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 tcaagaagaagcgaagcgaagcgaagccttctagatggaacaagcagaagat 120
OY 215 gtggaaggaaggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 gtagaaggaaggagcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
OY 275 tcaacgcgcgtgagcgaacatccccaagaacacaccccttgctcaagctgaccc 331

```

```

DB 181 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    tccaccgagcaggaagcaccatccccaaggccacaccccttgcccaagccgctcccttg 240
OY 332 gccggggtgggaccccaacaaaggtggagctgcctccctctgactgtatcacccca 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 gccgaggtgggacaaccccccaacgaagcggagacatccctccctctgactgtaccccca 300
OY 392 gctgagcctccagacagatgtagtgagcttgagccacggaattcccaagcagaagcc 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 gctccgactccaacacagaccatctgatacttgagcatalagattctcaagctcggcgcg 360
OY 452 tggaggtgtgggctagaagggcctctggaagaagagcctgcctcgtctgctcccgag 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 tgggaggtgagct---tgagctggtggaagaagaagccagcccgctgcctccccaag 417
OY 512 gccaccctcccaagctggtctggaatgacgaactcgcggaaccccgccagatccac 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 gtgtgttaccagcagcttggtggtgagatgtagctgtgcagaagccgggggcccaggtctac 477
OY 572 atgcgctcattgacgaagacacacctgtactagatgcatggaactagctccaagctagtc 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 atgcaactcattgacgaagacacacctgtactagatgcatggaactagctccaagctagtc 537
OY 632 atcttcgacaccatgctggagatcaagaagcctcttctgtctgtggtggccaagctgtg 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 atcttcgacaccatgctggagatcaagaagcctcttctgtctgtggtggccaagctgtg 597
OY 692 cgggcagccctctatggtggaacgaagaagcagcgttggtgggagctgaccatcaact 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 598 cgagcgggaccccttggtggagacgaagaagcagcgttggtgggagctgaccatcaact 657
OY 752 gacttcaactgtgtgtgtgacacgctactacaggtcccccctgtgtccagatctagat 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 658 gacttcaactgtgtgtgtgacacgctactacaggtcccccctgtgtccagatctagat 717
OY 812 gaacaacataagatgtgagacctggaaggagatcactcgaagcctgtctcaagcctct 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 718 gaagaacataagatgtgagacctggaaggagatcactcgaagcctgtctcaagcctct 777
OY 872 gtctcactctctcctaataagctgttgaagctgtctacacccctcatcaagaacgg 931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 778 gtctcactctctcctaataagctgttgaagctgtctacacccctcatcaagaacgg 837
OY 932 atccatgcgctgcgtgtctcttcttgaacgggtgtcaaggaacgtaactccatccatccacac 991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 838 atccacgcgtgcgtgtctcttcttgaacgggtgtcaaggaacgtaactccatccatccacac 897
OY 992 aaagcgtgtcgaagctcctgcgaacatcttgttctcctgtgtcccgccctctctctc 1051
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 898 aagcgtgtcctcgaagctcctgcgaacatcttgttctcctgtgtcccgccctctctctc 957
OY 1052 taaccgacatctcgaagatttgagcatcgcgcacatctcgaagcctgtgtgtgtgtgag 1111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 958 taaccgacatctcgaagatttgagcatcgcgcacatctcgaagcctgtgtgtgtgtgag 1017
OY 1112 acagcaccatctctgactgactgacatcgacatcttctgtgacccggtgtgtgtgtgtgag 1171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1018 acggtcccatctctgacatcgacatcttctgtgacccggtgtgtgtgtgtgtgag 1077
OY 1172 gtgtgtcaagaaatgtgtgagctgcgttggtggtctatctccgctcttgatgtgtgtgtgag 1231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1078 gtgtgtcaagaaatgtgtgagctgcgttggtggtctatctccgctcttgatgtgtgtgtgag 1137
OY 1232 gctgtcccaagaacacacacacacacacacacacacacacacacacacacacacacacac 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1138 gctgtcccaagaacacacacacacacacacacacacacacacacacacacacacacacac 1197
OY 1292 acactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1198 acactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1257
OY 1352 gacaggaattgtcgtggaggaaggtacacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


OY	758	atccggtgtcgtcatcgtctactacaaggtctcccccctgtgtccagatctatgaaattgaaaca	817
Db	658	atcttggtagcgacccgcatcttacaaggtctcccccctgtgtccagatctatgaaattgaa	717
OY	818	catagatcttgagaccttggaaggagatctactctgcaagaactcttcaaacctctgtctcc	877
Db	718	catagatcttggaaccttggaaggagatctactcttcaagaactcttcaaacctctgtctcc	777
OY	878	atctctctaatagtatgactgtcttctgaagcgtgtcttacaacccctcatcaagaaccgata	937
Db	778	atctctccaatgatgacgtctgtctgaagcgtgtcttcaagccctcatcaagaaccgata	837
OY	938	cgctcgctctgtcttcttgaccctggtgtctgaagcaacgtactcccatctcaacaacaagc	997
Db	838	cgctcgctcggtgtcttgaccctgtctcccggtgtgtctgtctcaacatctcaacaataagc	897
OY	998	ctgtctcaagctctcttcacatctctgtgtctccctgtctcccgccctctctctctacgc	1057
Db	898	cttctccaagctctcttcacatctctgtgtctccctgtctcccgccctctctctctacgc	957
OY	1058	atctacaagaatttgagcatctgagcaatctccgagaactgtgtctgtgtctggaagaca	1117
Db	958	atctacaagaatttgagcatctgagcaatctccgagaactgtgtctgtgtctggaagaca	1017
OY	1118	cccatctcggaactctgacactcttgacatctcttctgacccgctgtgtctgcaactctgtgc	1177
Db	1018	cccatctcggaactctgacactcttgacatctcttctgacccgctgtgtctgcaactctgtgc	1077
OY	1178	aacgaatgtgtctgaagctgt	1237
Db	1078	aacgaatctgt	1137
OY	1238	cagcaaaccttaaacaccactctgtgacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1297
Db	1138	caacaacaatataaacaccactctgtgacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1197
OY	1298	tgctcgagagagtgctctctctctctgacagcccccacgagaactctgtgtgtgtgtgtgt	1357
Db	1198	tgctcgagagagtgctctctctctctgacagcccccacgagaactctgtgtgtgtgtgtgt	1257
OY	1358	atctgtctcggtgagcatgtacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1417
Db	1258	atctgtctcggtgagcaacatgt	1317
OY	1418	gtgtgtctctctctctctcagacatctctcagagcaactgtgtgtgtgtgtgtgtgtgtgtgt	1477
Db	1318	gtgtgtgtctctctctctcagacatctctcagagcaactgtgtgtgtgtgtgtgtgtgtgtgt	1377
OY	1478	ctctggtgtctgtgaagaatctgt	1536
Db	1378	ctctggtgtgtctgtgaagaatctgt	1437
OY	1537	atgaagctgaagctgaagcaatctcagccttctctcccccacccatcttctgtctcagct	1596
Db	1438	atgaagctgaagctgaagcaatctcagccttctctcccccacccatcttctgtctcagct	1496
OY	1597	atgattcaggtgagctgt	1624
Db	1497	atgattcaggtgagctgt	1524
RESULT 7			
ID ABA08485 standard: cDNA: 547 BP.			
XX ABA08485;			
XX			
XX 11-JAN-2002 (first entry)			
XX			
XX Human AMP-activated protein kinase subunit homologue cDNA, SEQ ID NO:261.			
XX			
XX Human: cytokine; cell proliferation; cell differentiation; growth factor;			
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;			

XX inhibin;chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
KM atherosclerosis; coronary heart disease; arterial ischemia;
XX bone disorder; osteoporosis; vascular growth disorder;
KM tissue regeneration; wound healing; infection; immune disorder;
KM cell culture; drug screening; gene therapy; antiinflammatory;
KM antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KM cytosolic; osteoprolifer; vasotropic; cardiant; vitrucide; antibacterial;
XX antifungal; vulnerrary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PM MO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PM 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-457740/49.
XX
DR P-PSDB; ABB11241.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1: Page 429; 1963pp; English.
XX
XX Sequences ABB10981-ABR12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoietic regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.

AAC98774
 ID AAC98774 standard; cDNA; 1691 BP.
 AC AAC98774;
 DT 09-MAR-2001 (first entry)
 DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:2.
 XX
 KW Human: pancreas; pancreatic cancer; pancreatic cancer antigen; detection; diagnosis; identification; cytostatic; neuroprotective; neurotropic; immunomodulatory; relaxant; contraceptive; gynaecological; antiinflammatory; cardiant; gene therapy; chromosome mapping; linkage analysis; tissue identification; tissue typing; forensic; neural; immune system; muscular; reproductive; gastrointestinal; pulmonary; cardiovascular; renal; proliferative; ss.
 KW
 XX Homo sapiens.
 OS
 PN MO200055320-A1.
 XX
 PD 21-SEP-2000.
 XX
 PP 08-MAR-2000; 2000WO-US05989.
 XX
 PK 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM;
 DR WPI: 2000-579444/54.
 XX P-PSDB: AAB54009.
 XX
 PT New nucleic acid that is a pancreatic cancer antigen for preventing, treating, or ameliorating a medical condition, particular pancreatic cancer, or for use in assays for diagnosing a pathological condition -
 PT
 XX
 XX Claim 1: Page 498-499; 1379pp: English.
 XX
 XX AAC98773 to AAC99231 encode the human pancreatic cancer associated proteins, called pancreatic cancer antigens, given in AAB54008 to AAB54466. The human pancreatic cancer antigens have cytostatic, neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive, gynaecological, cardiant and antiinflammatory activities, and can be used in gene therapy. The polynucleotide and proteins can be used for preventing, treating, or ameliorating a medical condition or in assays for diagnosing a pathological condition or a susceptibility to one in a subject. Binding partners to the proteins and the activity of the proteins can be identified. The pancreatic cancer antigens can be used to detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The pancreatic cancer antigen polynucleotides can be used to design nucleic acid hybridization probes that can be used in chromosome mapping, linkage analysis, tissue identification and/or typing and a variety of forensic and diagnostic methods. The proteins can be used to generate antibodies which are used to purify, detect and target the polypeptides, including both in vivo and in vitro diagnostic and therapeutic methods. The proteins can be used to treat or prevent neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal or proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent sequences used in the exemplification of the present invention.
 CC
 XX
 XX Sequence 1691 BP; 458 A; 382 C; 401 G; 449 T; 1 other;
 SO

Query Match 23.9%; Score 393.4; DB 21; Length 1691;
 Best Local Similarity 65.1%; Pred. No. 1.2e-92;
 Matches 580; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

OY 567 tctacatcgcttcacagagacacccgtacatgcatgagcaactgctccaaac 626
 DB 123 tgcatactcctcctacgaagtcacatgctgcatgacctgattcccaagctccaat 182

OY 627 tagtcatcttcgaacacccatgctgagatccaagaagccctctctgctctggtgccaag 666
 DB 183 tggtttatttgatgaagctccctcaggtggaagaagctcttctgcttggtagtcaag 242
 OY 687 gtgtcgagcagccctctatggagcagcaagaagcagacttggggagatgctgacca 746
 DB 243 gtgtacagactgcccctttagttagtgaagaagcaaatcttctgtggcactgtgacca 302
 OY 747 tcaatgacttcatccctggtctcctacatgactacaaagctcccccctgtccagatcatg 806
 DB 303 tcaatgacttcatcaatataccctgacccgactacataataacgcttggtagcatctatg 362
 OY 807 agattgcacacataaagattgagacactgagagagatcactgcaagctgtccaaagc 866
 DB 363 agctagaagaacacaaagatagaacttggagagaggtatctccagactccctttaaac 422
 OY 867 cctcgtgctccatctctccctaaatgatatgctgtttagaagctgtctaaacctcatcaga 926
 DB 423 cgtctgtcctcattctccctaaatgacagctgtttagtctgtctcttcaatattcga 482
 OY 927 accgatccatcgctcgtcgtctctttagaccggtgtagcagcaagctacatccatccca 986
 DB 483 acaagatccacagagctccagttatgacccagaatcaggaatattgtatcatctcca 542
 OY 987 cacacaaagcctgctcaagttccctgacacatcttggctccctgctgcccgcctcct 1046
 DB 543 ccacaaagcagatctgaaatctcctcaaatgtttatcactgagttcccaagcagaagt 602
 OY 1047 tccctacgcacatcccaagattggcagatcgacatccgagaaacttggcgtgtgtc 1106
 DB 603 tcatgtccaaagctctcgtgaagatcagatctgacactatgccaatatgtcatgttcc 662
 OY 1107 tggagacagaccccatccctgactcactgacacatcttctggaaccggtgtgtctcac 1166
 DB 663 gcaatcaccaccccgctctatgctcgtctgggaatttcttgaagacatcggtccagccc 722
 OY 1167 tgcctgtggtcaacgaatgtgtcaggtcgttgagcctctatccgccttgatgtgattc 1226
 DB 723 tgcaggtgtgtgagaaagagggcgtgtgtgtgacatctcctcaaatgtgatgttca 782
 OY 1227 acctggtctcccaagcaaacactacacacacactgagatgtgtggaagcccttgagtc 1286
 DB 783 atctggcagcagaagaacactacacacacactagatgtatctgtgacaaagcccttgaa 842
 OY 1287 agaggaactatgtctggaagggagctcttctctgacagcccccagagagcttggggag 1346
 DB 843 atcgatcacattacttgaaggtgttctcaaggtgtctacactgacatgagacttgagacca 902
 OY 1347 tgatcgacagatgtctcgcggaagaggtatacagaagctggtgtgtgtagagagaccagc 1406
 DB 903 tcatcaacagcgtgtggaagcagaagttcacccgacttgaagtgtgtgagaaatgagt 962
 OY 1407 atctctggaggtgtgtctccctcgcagatccctcagcagctgtgctca 1457
 DB 963 tgcacaaaggagatgtatcatcactgtctgaacatccctcgaagccctgtgtctca 1013

RESULT 10
 AAT85927
 ID AAT85927 standard; DNA; 1576 BP.
 AC AAT85927;
 DT 16-MAR-1998 (first entry)
 DE Mammalian AMPK-gamma subunit DNA.
 XX
 XX
 KW 5'-AMP activated protein kinase; AMPK: non-catalytic subunit; protein phosphorylation; cholesterol; fatty acid; gamma subunit; hormone sensitive lipase; HSL; ss.
 KW
 OS Mammalia.

PT	full-length cDNAs -
XX	
PS	Claim 8; SEQ ID 12660; 2537pp + CD ROM; English

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-*dT* primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Query Match	22.3%	Score 366.6;	DB 22;	Length 2223;
Best Local Similarity	62.98;	Pred. No. 1.3e-85;		
Matches 567; Conservative	0;	Mismatches 334;	Indels 0;	Gaps 0;

[illegible][illegible]

RESULT 13

ID ABL18857 standard; DNA; 3261 BP.

AC ABL18857,

DT 26-MAR-2002 (first entry)

DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 8044
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 8044

..... KW Drosophila; developmental biology; cell signaling; insecticide;

KW pharmaceutical; gene; ds

Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P

XX
XX
DEFE , DE COND WY[illegible]

XX WPT: 2001-666960/75
DB

XX
XX
XX

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions.

XX	
PS	Claim 1: SEO TD NO R044: 21no + Sequence listing: English

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB161516-AB130511), expressed DNA
CC sequences (AB101840-AB161515) and the encoded proteins
CC (AB557737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from *WIPO*
CC at [ftp://wipo.int/pub/published](http://wipo.int/pub/published) pct sequences.

PR 08-NOV-2000: 2000US-0246528.
PR 08-NOV-2000: 2000US-0246532.
PR 08-NOV-2000: 2000US-0246609.
PR 08-NOV-2000: 2000US-0246610.
PR 08-NOV-2000: 2000US-0246611.
PR 08-NOV-2000: 2000US-0246613.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249264.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249297.
PR 17-NOV-2000: 2000US-0249299.
PR 01-DEC-2000: 2000US-0250160.
PR 01-DEC-2000: 2000US-0250160.
PR 05-DEC-2000: 2000US-0250391.
PR 05-DEC-2000: 2000US-0251030.
PR 05-DEC-2000: 2000US-0251088.
PR 05-DEC-2000: 2000US-0251719.
PR 06-DEC-2000: 2000US-0251479.
PR 08-DEC-2000: 2000US-0251856.
PR 08-DEC-2000: 2000US-0251868.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251989.
PR 08-DEC-2000: 2000US-0251990.
PR 11-DEC-2000: 2000US-0254097.
PR 05-JAN-2001: 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
P1 Kosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS
PS Disclosure: SEQ ID NO 27552; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAK91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 3425 BP: 916 A; 794 C; 951 G; 764 T; 0 other;

Query Match 15.38; Score 252; DB 22; Length 3425;

Best Local Similarity 100.08; Pred. NO. 1.3e-55;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1396 cgaagccagacatctcttggcgigtctccctcgcagacatcctcagcactggtgt 1455
|||||
Db 3425 CGAGACCCAGCATCTCTTGGGCGTGTCTCCCTCTCCGACATCTTCCAGGCACTGTGCT 3366
Qy 1456 cagccctgtgcatcgaatgcctcctgggcttgagaagatctgagctccaatcccaagc 1515
|||||
Db 3365 CAGCCCTGCTGCATCGATGCCCTTCGGGCGCTGAGAAAGATCTGAGTCTCAATCCCAAGC 3306
Qy 1516 cactgcacacctggaagccaatgaaggaaactggagaactcagcctcaatctcccca 1575
|||||
Db 3305 CACCTGCACACCTGGAAAGCAATGAAAGGAACCTGAGAACTCAGCCTTCATCTTCCCCA 3246
Qy 1576 ccccatctgtgttcagctatgatcaggtaagctctgccttgggcatgacacagc 1635
|||||
Db 3245 CCCCATTGCTGTTTCAGCTATGATTTAGGTAGGCTTGCTTGCCCATGACACAGC 3186
Qy 1636 ctcttagctctc 1647
|||||
Db 3185 CTCTTAGTCTTC 3174

Search completed: October 3, 2002, 16:30:41
Job time: 14311 sec


```

Db      127  TGGTTGATTATTGATFACCTCCCTGCAGCTGAGAAAGCTTTTTCCTTGGTACTTAACG 186
Oy      687  gtlgcygagcagccctctctatggagacagaagaagagcttltgltggagatgtaccca 746
Db      187  GTGTAGAGCTGCCCTTTATGGATAGTAGAAGCAAGTTTGTGGGATCTGACCA 246
Oy      747  tcaatctatctaccggtgtgtacgtctacacaggttccccctgtgtccagatctatg 806
Db      247  TCACGTATTTCATCAATATCTGACCGCTACTATTAATCAGCCTTGGAACATCATATG 306
Oy      807  agatltgaacacataagatltgagaccttgagagagatctacacagagctgtctcaagc 866
Db      307  ACCTACAGACAGCAAGATAGAAACTTGGAGACAGCTGTATCTCCAGACTCTTTAAAC 366
Oy      867  ctctgtgtccatctctcctaaatgataagcctgtlttgaagctgtctacacccatcaaga 926
Db      367  CGCTTCTCTGATCTCTCTTAATGCGCAGCTTGTGTATCTGCTCTCTTATTAATTCGGA 426
Oy      927  accggaatcagcctgtgctgtctgtaccggtgtcaggaagcgtacacatctctca 986
Db      427  ACAGATTCACAGGCTGCGCAGTTATTTGACCCAGAAATCAGCAATACCTTTGTATCTCA 486
Oy      987  cacacaaagcctgtcctcaagtlcctgtcacatctltgttccctgtgtcccgccctct 1046
Db      487  CCCACAAAGCGCATTTGTGAAGCTTCTCAATTTGTTATTCATCTAGTTCCCAAGCCAGAT 546
Oy      1047  tcccttaccgcacatctcaagatlttgagacatcgcgcacatctcgagactgtgtgtgac 1106
Db      547  TCATGTCCCAAGTCTCTGGAAGACTACAGATTCGCACTATGCTCAATATGCTATGCTTC 606
Oy      1107  tggagacagacacccatctccgactgtcagctgtgagatctlttggacggggtgtgtctgac 1166
Db      607  GCACATACACCCCGCTTATGTGCTCTGTGGGATTTTGTACAGCATCAGATCTAGCCCC 666
Oy      1167  tgcctgtgtcagaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1226
Db      667  TCCAGTGTGCTGATGAGAAAGCGCGCTGTGCTGATCATCTACTCAAGTTGATGTATCA 726
Oy      1227  accgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1286
Db      727  ATCTGTGCAACAGAAAGACCTTACACACCTTATGATGTATCTGTGACTTAAAGCTTTGCAAC 786
Oy      1287  agagacacatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1346
Db      787  ATCGATTCACATTTAGTGTAGGCTGTCTCAAGTCTGCTACCTGATGATAGACTCTGGAGACA 846
Oy      1347  tgaatgacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1400
Db      847  TCATTCACAGGCTAGTGTGAGACGACAGCTTTCACCGCACTTGTGCTGTGTGATGAA 900

```

RESULT 6
AL555228
LOCUS AL555228 905 bp mRNA linear EST 16-FEB-2001
DEFINITION AL555228 L/T1_NFL006_P12 Homo sapiens cDNA clone CS0DK007YE02 5 prime, mRNA sequence.
ACCESSION AL555228
VERSION AL555228.1 GI:12896759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 905)
AUTHORS L.L.M.B., Gruber,C., Jessée,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
 source 1..905

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK007YE02"
/clone_1db="L/T1_NFL006_P12"
/tissue-type="placenta"
/notes="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 237 a 218 c 199 g 250 t 1 others
ORIGIN
Query Match 22.1%; Score 363.2; DB 9; Length 905;
Best Local Similarity 64.9%; Pred. No. 5.2e-75;
Matches 536; Conservative 1; Mismatches 289; Indels 0; Gaps 0;
Oy      567  tctacatgcgtctcatcagagagacacacctgtacatgtccatgtggaactagctcacaagc 626
Db      80  TGTATATCTTCCCTTCATGAAAGTCTCATCGCTGATGACCTGATTCACAAAGCTCAAT 139
Oy      627  taatcatcttgcagcccatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 686
Db      140  TGGTTGTATTGATATGTCCTCGACGGTGAAGAAAGCTTTTTCCTTTGCTGTGATCAACG 199
Oy      687  gtlgcygagcagccctctctatggagacagaagaagagcttltgltggagatgtaccca 746
Db      200  GTGTAGAGCTGCCCTTTATGGATAGTAGAAGCAAGTTTGTGGGATCTGACCA 259
Oy      747  tcaatctatctaccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 806
Db      260  TCACGATTTTATCAATATTCCTGCACCGCTACTATTAATTCAGCCTTGTGTACAGATCTATG 319
Oy      807  agatltgaacacataagatltgagaccttgagagagatctacacagagctgtgtgtgtgt 866
Db      320  AGCTAGAGAAACACAGATTAAGAACTTGGAGACAGGCTGTATCTCAGAGACTCTTTAAAC 379
Oy      867  ctctgtgtccatctctctctaaatgataagcctgtlttgaagctgtgtgtgtgtgtgtgt 926
Db      380  CGCTTGTCTGCAATTTCTCTTAATCCAGCTTGTGTGATGCTGTCTCTTCAATTAATTCGGA 439
Oy      927  accggaatcagcctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
Db      440  ACAAGATTCACAGGCTGCGCAGTTATTTGACCCAGAAATACAGCAATATCTTGTATCTCTCA 499
Oy      987  cacacaaagcctgtcctcaagtlcctgtcacatctltgttccctgtgtcccgccctct 1046
Db      500  CCCACAAAGCGCATTTGTGAAGCTTCTCAAAATTTGTTATTCACATGAGATTCACCAAGCAAGT 559
Oy      1047  tcccttaccgcacatctcaagatlttgagacatcgcgcacatctcgagactgtgtgtgtgt 1106
Db      560  TCATGTCCCAAGTCTCTGGAAGCTTACAGATTTGGCACCTATGCAATATATGCTATGCTTC 619
Oy      1107  tggagacagacacccatctcagctgtgagacatctlttggagacggggtgtgtgtgtgtgt 1166
Db      620  GCATATACACCCCGCTTATGTGCTCTGTGGGATTTTGTGACAGCATGAGTCTCACACCC 679
Oy      1167  tgcctgtgtcagaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1226
Db      680  TGCACGTGTGTGATGAGAAAGCGCGTGTGTGACATCTCAAGTTTGTGATGATATCA 739
Oy      1227  accgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1286
Db      740  ATCTGTGAGAGCAAGAAAGCTTACACACCTTATGATGTGACTTAAAGCTTGTGCAAC 799
Oy      1287  agagacacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1346

```

Dd	800	ATCGATCACATTACTTTGAGGGGTCTTCAAATGCTACCTGCATGACACTTGAGAGACA	859
Oy	1347	Tgaatgacagaaatctcggagacgagccaggctgtgcatact 1392	
Dd	860	TTCAATCAACAGCCTACTGTCGACAGAGCTTACCAGCTTGTACTGTGT 905	
RESULT	7		
AL554278			
LOCUS	AL554278 LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1082YL23 5	940 bp mRNA linear EST 16-FEB-2001	
DEFINITION	prime, mRNA sequence.		
ACCESSION	AL554278		
VERSION	AL554278.1 GI:12894901		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES	location/Qualifiers		
source	1..940		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0D1082YL23"		
	/clone_1id="LTI.NFL006.PL2"		
	/tissue_type="placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by lile technologies. Contact : Feng Liang lile technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifestech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	248 a 222 c 208 g 260 t	2 others	
ORIGIN			
Query Match	21.9% Score 361; DB 9; Length 940;		
Best Local Similarity	64.8%; Pred. No. 1.8e-74;		
Matches 535; Conservative	0; Mismatches 250; Indels 0; Gaps 0;		
Oy	567	tctacaatgcgttcataatgcagagcacaccctctcaatgatcatatgcagaatcagc 626	
Dd	93	TGTATTAATCTCCATCATGAAGCTCATGCCCTCTATGACCTCATTCGCCAAGCTCCAAT 152	
Oy	627	tagtcatctcgacaaccaatgccttgagatcaagaagagcccttcgtcctgtgtgccaa 686	
Dd	153	TGGTGTATTTGATGATACGTCCTCGACAGTGAAGAAGCTTTTTTGTGCTTGGTGAATAACG 212	
Oy	687	gtctgcgggcagcccctctataggacaagcaagaagcagctttgtgggagtgtgacca 746	
Dd	213	GTTGTACAGAGCTGCCCTTTATAGCATAGTAGAAGCAAAGTTTGTGGAGTGCATGACCA 272	
Oy	747	tcaactgacctcaatccgtgtcgtcatcgatcgctactaacaggtccccctgtccaatctat 806	
Dd	273	TCACTGATTTTCATCATATATCTCTGCACCGCTTACTATTAATATCAGCCTTGATACAGATCAT 332	
Oy	807	agattgacaacaabaagattgagacccttggagagatctcaactgtcaaagctcctaagc 866	
Dd	333	AGCTAGAAGAACACAAAGATAGAAACTTGGAGAGAGGCTATCTCCAGAGACTCCTTAAAC 392	
Oy	867	cctcgtgctccatctctcccttaatgatagatgcctglttgtgaagctgtgtctacaacctcaaca 926	

FEATURES	LOCATION/Qualifiers
Seq primer: ATTGAGGTGACACTATAG.	Row: 1 Column: 16
PCR primers	
and -mismatch 12 options.	
Single pass sequencing. Bases called and alt_trimmed with phred	
VO.980904.e. Vector identified by cross_match with the -mnscore 18	
FORWARD: AGGAACAGCTATGACCAT	
BACKWARD: GTTTCCCGACGTACACGAC	
Plate: 90	
Row: 1	
Column: 16	
Seq primer: ATTGAGGTGACACTATAG.	
Location/Qualifiers	


```
Db 731 TCCACGTGCTGATGACGAAGCGCGCTGTGTCACATCTACTCCAAAGTTTGATGTTATCA 790
Oy 1227 accgagcgtcccaagcaaaccttacaacacclgagacatgagtgaggagagccctgagac 1286
    1 1111 11 1 11111111 1111 11 1111 111111 11
Db 791 ATCTGGCAGCAAGAAAGACCTACACACACTAGATGATCTGTGACTAAAGCCTTGCAAC 850
Oy 1287 agagagccatcgtctgagaggagctccttcctgcgaagcccaagagagcttgagg 1342
    1 1111 11 1 111111 1111 1111 11 111111 1111:1
Db 851 ATCGATCACATTACTTGAAGGCTGTTCACAGCTCTACCTGCATGAGACTCTGGK 906

RESULT 10
LOCUS B1858240 864 bp mRNA linear EST 10-OCT-2001
DEFINITION 603384001P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5392777 5',
            mRNA sequence.
ACCESSION B1858240
VERSION B1858240.1 GI:15998987
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 864)
            NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-femail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LML at:
            http://image.lnl.gov
            Plate: LLM12001 row: a column: 02
            High quality sequence stop: 830.
            Location/Qualifiers
                source
                    1..864
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5392777"
                        /clone_lib="NIH_MGC_87"
                        /isue_type="mammalian adenocarcinoma, cell line"
                        /lab_host="DH10B (phage-resistant)"
                        /note="Organ: Breast; Vector: pCMV-SPORT6; Site_1: NotI;
                        Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
                        Average insert size 1.383 kb. Library enriched for
                        full-length clones and constructed by Life Technologies.
                        Note: this is a NIH_MGC Library."
BASE COUNT 223 a 202 c 203 g 236 t
ORIGIN
```

```
Db 182 GCCTTGTAGACATCTATGACCTAGAGAAACACAAAGATAGAAACTTGGAGAGCTGAT 241
Oy 848 ctgcaagcgtcttcaagccctcgtgctccalcctccctaaatgataagcgtgttgaagc 907
    1 1111 11 11 111111 1111 1111 11111111 11111111 11
Db 242 CTCACGACCTCTTAAACCCGCTTGCTGACTTCTCTTAATGCCACCTTTGATGCT 301
Oy 908 gtctacaacccctcaagaagaccggaatccatcgtctgctgtctcttgaccgggtctggc 967
    1111 11 11 1111 111111 111111 111111 111111 111111
Db 302 GTCCTTCTATTAAATTCGAAACAAAGATCCACAGGCTGCAGTATTGACCCACGAATCAGG 361
Oy 968 aacgtactccacatccctcacacaacaagccctgtcgaagcttcgcacatccttggctc 1027
    11 11111111 111111 11111111 11111111 11 1111
Db 362 AATACTTTGTACATCTCCTCACCCACAAAGCCATTCTGAAAGTTCCCAAAATTTTATAC 421
Oy 1028 ctgtgcgcccgccctccctccctccctccctccctccctccctccctccctccctcc 1087
    1 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 422 GAGTTCCCAAGCAGAGATTCATGCTCCAAAGCTCTGGAAAGAGCTACAGATTGGCACC 481
Oy 1088 cgaagactggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1147
    1 1111 1111 11 1111 11 1111 11 1111 11 1111 11 1111
Db 482 GCCAATATTGCTATGTTGCTGCACCTACCAACCCGCTCTATGCTGCGGATTTTGTGA 541
Oy 1148 gaccggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1207
    1 1111 11 1111 111111 1111 11 11 111111 1111
Db 542 CAGCATCGACGTCTCAGCCCTCCAGTGTGTGATGAGAAAGGCGCTGTGTGACATCTAC 601
Oy 1208 tcccgcttgatgtgattcaccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1267
    111 11111111 11 1111 11 1111 11 111111 1111 11
Db 602 TCCAAATTTGATGTTATCAATCAATCTGGACAGCAAAAGACCTACAAACATGATGTAT 661
Oy 1268 gtggaggaagcccgagagcgagagacactaagctgtgtgtgtgtgtgtgtgtgtgtgt 1327
    111 1111 1111 11 11 1111 11 1111 11 1111 11 1111
Db 662 GTGACTAAAGCCTTGCAACATCATGATCACTTCTTAGGGGTGT-CTCAAGTGTCTACTG 720
Oy 1328 caacgaagcgtggggaagtgatcagacagatgt-cctcgagagcaggtacacgaagcgt 1386
    1111 1111 1111 11 1111 111111 11 11 1111 1111 11
Db 721 CATGAGACTCTCGAGACCATCATCAACAGCGTACTGCAAGCCGAGGTTCCACCATTTGT 780
Oy 1387 gctagtgaagacgaacccaagcatcctctgtggcgtgtgtgtgtgtgtgtgtgtgtgtgt 1446
    1 1111 11 1111 11 1111 11 1111 11 111111 1111
Db 781 AGTGCTGATGATAAATGATGTGTGTCCAGGAGCATGT-TCACCTGTGTGACATCTCGAGG 839
Oy 1447 actgtgtgtc 1456
    11111111
Db 840 CTGTGTGCTC 849

RESULT 11
LOCUS AL519198 826 bp mRNA linear EST 13-FEB-2001
DEFINITION AL519198 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA0127B06 5
            prime, mRNA sequence.
ACCESSION AL519198
VERSION AL519198.1 GI:12782691
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 826)
            L.I.W.B., Gruber,C., Jesse,J., and Polayes,D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
JOURNAL Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
            source
                1..826
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="CS0DA0127B06"
                    /clone_lib="LTI_NFL011_NBC1"
```



```

/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DHI10g"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	20.88;	Score 343;	DB 9;	Length 826;
Best Local Similarity	65.78;	Pred. No. 2.9e-70;		
Matches 499; Conservative	0;	Mismatches 260;	Indels 0;	Gaps 0

DY 567 tctatcatgagcgtcaatgcagagacacccctgtcacgaatgcaactgatctccaagc 626
+
Db 68 TGTATACCTTCTTCATGAGCTCATCGCCTGTATGACTGATTCCACAAGCTCCAAT 127

Dy 627 tatcatccttcgaaccatgcttgagaatcaagaagccctcttgcctgtgccaacg 686
| | | | | | | | | | | | | | | | | |
Db 128 tgcgttatattgatcgctccctgcagctgaagaaagcttttttgccttggctgactaacg 187

Qy 687 gctctgcgagcagccccctctatrgygacagcaagaagccttctgtggygatgtcacca 746
|||||
Db 188 cgtctacgacctgccccctttatgggattagtaagaacaaagttttctggcgatgcctaccca 247

Oy 747 tcaactgacttcacccctgbcctgcatcgctactacaggtcccccctgtccagatctatg 806
||||| ||||| | ||||| ||||| | | ||||| ||||| |||||
Db 248 tcaactgattttcatcatatattcctgcaccgcctactataaattcagcccttggtacagatctatg 307

| | | |
|----|--|-----|
| Oy | 807 agatltgaacaacataagatctgagaccigaggagagatcttacctgaagccttgctccaagc | 866 |
| | | |
| Dd | 308 acgttagaacacacacagattgcmaactttggacagagagtatatcttcacagaactccctttaac | 367 |

Oy 867 cttctgtctccatcctctcctaatagaagcctgtttagaacgtgcttaacacctcaatcaaga 926
| | | | | | | | | | | | | | | | | | |
Db 368 CGCTGTGTCGACATTTCTCCCAATGCAGCGTGGTGGTAAGCAGTGCGTGGTAAATTAAATGCCA 427

[illegible]

987 cacacaacgcctgctcaagttctctgcacacattgttccctgctgccccgcctctct 104

1047 tccctaccgcactatccaagattggcgatcgcacattccgagactgtgtgtgc 110

QY 1107 tggagacagcaaccatccctgactgacactggaacatctcttctggaacggcgctgtgtctgcac 116

1167 tgcctggtcaacgaatcgtgltcaagtcgtgggcctctatccgccttgatgtgattc 122

1227 acctgctgcccagcaaacctacaacccaactgacatgagtctggagagaagccctgagc 128

DB 128 ATCTGGCAGCAGAAAAGACCTACACCAACCTAGATGTATCTGTGACTAAAGCCTTGCAC 787

QY 1287 agaagcaactatgtctggaaggagtccttctcgcagc 1325

Db 788 ATCGATCACATTACTTTGAGGGTCTCTCAAGTGCTACC 826

| | RESULT | 12 |
|------------|--|------------------------------------|
| LOCUS | AL513784 | |
| DEFINITION | AL513784 LTI_NFL006_P12 Homo sapiens cDNA clone Cl08A0032H07 5 prime, mRNA sequence. | 951 bp mRNA linear EST 13-FEB-2001 |
| ACCESSION | AL513784 | |

| | |
|----------|--------------|
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |

REFERENCE 1 (bases 1 to 951)

| | |
|---------|--|
| TITLE | Full-length cDNA libraries and normalization |
| JOURNAL | Unpublished (2001) |
| COMMENT | Contact: Genoscope |

FEATURES

BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Location/Qualifiers

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. 951 |

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CLOBA003ZH07"
/clone_1lb="L11-NFL006_PL2"
/tissue_type="Placenta"

```

/note=Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA fax: (1) 301 601 8371

| | | | | | |
|------------|---|-------|-------|-------|----------|
| BASE COUNT | 295 a | 190 c | 213 g | 252 t | 1 others |
| ORIGIN | email : lianqing@lianecon.com URL :
http://fulllength.invtrogen.com" | | | | |

| Query Match | Score | DB | Length |
|-------------|--------|----|--------|
| 20.88; | 342.6; | 9; | 951; |

| | | | | | | | | | |
|---------|------|--------------|----|------------|------|--------|----|------|----|
| Matches | 533; | Conservative | 0; | Mismatches | 317; | Indels | 0; | Gaps | 0; |
|---------|------|--------------|----|------------|------|--------|----|------|----|

QY 567 tctacatgcgcttcata^gcag^gagcaca^cct^gctac^gat^gccat^ggcac^tta^gctcca^gc 626

Db 101 TTACATGCGATTTCATGAGGTCACACAAGTGTATGACATCGTCCAACCAAGTTCAAAGC 160

QY 627 tagtcattctcgaccacatgctgagatcaagaagccttcttctgtctgtgccaacg 686

Db 161 TTGTTGCTTTGATACTACATTACAGTTAAAAAGCCTTCTTTCCTTTGGTAGCCAACG 220

QY 687 gtgtgcggcagccctctatgagacagcaagaagcagaactttgttgqgatatqctgacca 746

Db 221 GTGTCGAGCAGCGCCACTGTGGAGAGTAAAAAACCAAGTTTGTAGGAATGCTAACAA 280

QY 747 tcaactgactcatcctgtgtcgcatacgtactactacaggtccccctgtgccagatctatg 806

Db 281 TTACAGATTTCATAAATATACTACATAGATTAATAATCACCCTATGGTACAGATTATG 340

QY 807 agattgaacacataagatttgagacctggaggagatctactctgcaagctgtctcaagc 866

Db 341 AATTAGAGACATAAAATTGAACATGGAGGAGCTTTATTACAGAACAATTAGC 400

QY 867 ctctggtctccatctctcctaatagtatagcctggttgaaagctgtctacacccctcatcaaga 926

Db 401 CTTAGTGAATATCTCCAGATGCAAGCCTCTTCGATGCTGTATACTCCTTGATCAAAA 460

QY 927 accggaatccatgcgcctgttctctgaaccggtgtcagagcaacgtactccaatcctca 986

```

Db 461 ATAAAAATCCAGATGCCCCGTTATTGACCCATATCAGTGGGAATGCACCTTTATATACTTA 520
Oy 987 caccacaacgctgctcgaagcttcctgcacatcttggttcctcgtgccgcgcctcct 1046
Db 521 CCCACAAAAGATCTCTCAAGTCTCTCAGCTTTTATGTCTATATATGCAAAAGCCCTGCT 580
Oy 1047 tccctaccgcacatcacaagatlttgagcaltgcgacatccgagacttgctgtgtgc 1106
Db 581 TCATGAACGAGAACCTGATGAGCTTGGAATAGGAACGACCAACAACATTCGCTTCATAC 640
Oy 1107 tggagacagaccacatccctgactgcacgtgacatcttctgtgagacgggtgtgtgcac 1166
Db 641 ATCCAGACACTCCCATCATCAAAAGCCTTGACATATTTGTGAAAACCAATATACAGTC 700
Oy 1167 tgcctgtgtcaacgaatgtgtcagtgctgtgagcctcctatccgccttgatgtgac 1226
Db 701 TCGCTCTTGTGATGATGACGAAAGTTGTAGATTTATTTATTCCAAATTTGATGTAATTA 760
Oy 1227 acctgtgtcccaagaacactacaacaccctgagacatgagtgtggagaagccctgagac 1286
Db 761 ATCTTCTCTCTGAGAAACATACAAATACCTAGATATTCACGCGTGACCCGACGCTTCAC 820
Oy 1287 agagagacatagtctgagggagagctccttcctgcagaccacagagagcttggggagag 1346
Db 821 ACCGTTACAGATTTTGAAGGTGTGTGAAGTGCATAATAGCTGGAATACTGGAGACCA 880
Oy 1347 tgaatcagagatgtctgagagcaggtacacagagctgtgtgtgtgtgagagaccagc 1406
Db 881 TCCTGCACGAATATGTAAGAGCTGATGCTCATTCGCTGTGTGTGTAATTAAGAGAMA 940
Oy 1407 atctctgtgg 1416
Db 941 GTATTGTGGG 950

RESULT 13
BM462694 1013 bp mRNA linear EST 05-FEB-2002
LOCUS BM462694
DEFINITION AGENCOURT_6427350 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5540312
5', mRNA sequence.
ACCESSION BM462694
VERSION BM462694.1 GI:18511734
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1013)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12236 row: d column: 09
High quality sequence stop: 640.
Location/Qualifiers
1..1013
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5540312"
/clone_1fb="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectional; Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life

```

```

BASE COUNT 272 a 241 c 224 g 275 t 1 others
ORIGIN
Query Match 20.4%. Score 336.6; DB 10; Length 1013;
Best Local Similarity 65.5%; Pred. No. 1e-68;
Matches 492; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

Oy 567 tctacatgcctcattcagagagacacacctgcatacatgacatgacatgacatgacacg 626
Db 70 TGTATACCTTCTTCAATGAAGTCTCATCCGCTGTATGACCTGATCCCAAGCTCCAAAT 129
Oy 627 taatcattcagacacatcgtcgtgagatcaagaagcctcctctgtcctgtgtgacag 686
Db 130 TGTGTGATTTGATAGTCGCCCTGACAGTGAAAGCAAGCTTTTGTGCTTGTGATGAACG 189
Oy 687 gttgtgagcagccctctatagagacagaagaacagagcttlttgaggatgtgacga 746
Db 190 GTATGAGAGCTCCCTTATAGGATAGTAAAGCAAGATTTTGTGGCATGCTGACCA 249
Oy 747 tcaatgactcaatccctgtgtgtcagtcagtaacaggtcccccctgtgtccagatcag 806
Db 250 TCAGTATTTTCAATATATCTGACACCGCTACTATAAATCAGCCTTGATACAGATCTATG 309
Oy 807 agattgaacaacataagattgagacctgtgagggagagatcactcgtcaaggtcgtcagac 866
Db 310 ACCTGAGAGAACACAAATAGAAACCTTGAGAGAGCTATCTCCAGAGATCTTTAAAC 369
Oy 867 ctctgtgtccatctctccttaagtagccgtgtgtgaagcgtgtcagacccctcacaag 926
Db 370 CGCTGTCTGCAATTTCTCTTAATGCCAGCTTGTGTGATGCTGTCTCTTCAATTAATCG 429
Oy 927 accgagatccatgcctgtcctgtcctgtgacccgtgtcagcaagctactcacaatccca 986
Db 430 ACAAGATCCACAGCGCTGCCAGTATTGACCCAGATCAGCAATCTTTGATACATCTCA 489
Oy 987 caccacaacgctgtcctgaagcttcctgcacatcttctgtcctcgtgtgccgcgcctcct 1046
Db 490 CCCACAAAGCGCAATCTGAAAGTCTCAAAATGTTATACAGAGTCCCAAGCCAGAGT 549
Oy 1047 tccctaccgcacatcacaagatlttgagcaltgcgacatccgagagactgtgtgtgc 1106
Db 550 TCATGTCCAAAGTCTCTGGAAGAGCTACAGATTCGACCATTCCAATATGTCTATGTTTC 609
Oy 1107 tggagacagaccacatcctcgtacatgcacatgacatcttctgtgagccgggtgtgtcac 1166
Db 610 GCACATACCAACCCCGCTTATGTGCTGTGGGATTTTGTACAGCATCGAGTCTAGGCC 669
Oy 1167 tgcctgtgtcaacgaatgtgtcagtgctgtgagcctcctatccgccttgatgtgac 1226
Db 670 TCCAGATGTTGATGAGAAAGGCGCTGTGTGACATCTACTCCAAAGTTGATGTTATTA 729
Oy 1227 acctgtgtcccaagaacactacaacacttgagacatgagtgtggagaagccctgagac 1286
Db 730 ATCTGCGACGAGAAAGAAAGCTCAACAACCTGATGATCTGTGATAAAGCTTGCAAC 789
Oy 1287 agagagacatagtctgtgagggagagcttct 1317
Db 790 ATCGATCACATTACTTTGAGGGGTGTCTC 820

RESULT 14
BM470188 1004 bp mRNA linear EST 05-FEB-2002
LOCUS BM470188
DEFINITION AGENCOURT_6476961 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553194
5', mRNA sequence.
ACCESSION BM470188
VERSION BM470188.1 GI:18519230
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


QY 567 tctacatgcgcttcacatgcagagcacacacctgtctacgatgcatggaactagctccaagc 626

| | | | |
|----|------|--|------|
| OY | 987 | cacccaagacgcctgctcaagllctctgacacclcttggctccctgctgcccgccctc | 1046 |
| | | | |
| Db | 603 | CCCAACAAAGATCTCTCAAGCTCTCTCCAGCTTTTATGTCGTATATGCCAAAGCGCTGCT | 662 |
| OY | 1047 | lctclacccgcacatclacaaagalltggacatcgccacatccgaagacttgcgtgtggtgc | 1106 |
| Db | 663 | TCTATGAAGCAGAACCTGGATGGACTTGGAAATAGCAAGCTACCAACAACATTCGCTTCATAC | 722 |
| OY | 1107 | ctggagacagcagcccaatccctgaactgcacatggacalcttgtgaccgagtggtgtgac | 1166 |
| Db | 723 | ATCCAGACACTCCCATCATCAATCAAAAGCCCTTGACACATATTTCTGAAAGACGAATATACGCTC | 782 |
| OY | 1167 | tgacctgtgtcaacgaatgtgtgtcagtgctgtagccctcatcccgcttatatgacac | 1226 |
| Db | 783 | TGCCGTGTGGATGAGACTCAGGAAGCTGTAGATTTATTATTCACAAATTTGATGATTA | 842 |
| OY | 1227 | accgtgctgcccaagaaacctacacaccctgacatgagtggtggagaagaccttgagc | 1286 |
| Db | 843 | ATCTTCCTCTCGAAGAAACATATACAAATACCTAGATATCAGCGTGACCCAGGCCCTTCAGC | 902 |
| OY | 1287 | agaagacacatctgtctggaaggaagtccttccctgcagcagcccaagagccttgggggaag | 1346 |
| Db | 903 | ACGCTTCACAGTATTTTGAAGGCTTTGTGAAGTGCATATAGCTGGAAATTAAGTAGAGACA | 962 |
| OY | 1347 | tgatcagcaggaatgtctcgggaagcagtaacaaagctggtgtctagtggacagaaaccagc | 1406 |
| Db | 963 | TCTGTGACACAAATAGTAAAGAGCTGAGGCTCATTCGGCTGTGTGTAAATAGACACAGTA | 1022 |
| OY | 1407 | atctcttgccgctggtctccctctcagacatccttcagagcaatgtgtgtcagccctgtg | 1466 |
| Db | 1023 | GTTATTTGGGTATTAATTTTCCCTCTCTCGGACATTTCTCAAGCCCTGATCTCTACACCCACAG | 1082 |
| OY | 1467 | g 1467 | |
| Db | 1083 | G 1083 | |

RESULT 3
 US-09-272-796-14
 Sequence 14, Application US/09272796
 Patent No. 6207148
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Corley, Neil C.
 APPLICANT: Guégler, Karl G.
 APPLICANT: Lal, Preeti
 APPLICANT: Goli, Surya K.
 APPLICANT: Shah, Purni
 TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
 NUMBER OF INVENTIONS: KINASES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/272,796
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/878,989
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J J

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PENITUT01
CLONE: 1452972
US-09-272-796-14

| | | | | |
|---------------------------|-------|--------------------|-----------|--------------|
| Query Match | 22.3% | Score 366.6; | DB 4; | Length 1435; |
| Best Local Similarity | 62.9% | Pred. No. 2.8e-87; | | |
| Matches 567; Conservative | 0; | Mismatches 334; | Indels 0; | Gaps 0; |

| | | | |
|----|------|--|------|
| QY | 567 | tcacatgcgcgtctctatcgaagagacaccctgcgtacagtcgcacatgcgaactacgtccaaagc | 628 |
| Db | 183 | TTTACATGCGGATTATGAGGTCACACACAGTGTATGACATCGTTCCAAACCAAGTTCCAAAG | 242 |
| QY | 627 | taagcatcttcgaacacatgctgtagataccaagaagcctctctctgctctgtagtgccaaag | 686 |
| Db | 243 | TTGTGTCTTGATTACTACATACATTAACAGTTTAAAAAGGCTCTTGTGCTTTGGTGGCAAG | 302 |
| QY | 687 | gttgcgcgagagcccccctctatgtagacgaagaagcagagctcttgtagggatagtcgaccca | 746 |
| Db | 303 | GTCGCCAGCAGGCCCTACTGTGGAGAGTAAAAAACAATTTTGTAGGAATCTTAACAA | 362 |
| QY | 747 | tcactgaacttcctctgtagtcgcatgcgtacagatcagagtcgcccccctggtccaaatcatg | 806 |
| Db | 363 | TTACAGATTTTCATTAATATCTACATAGATAGTATTAATCACCTATGCGTACACATTTATG | 422 |
| QY | 807 | agatctgaacacataagatgtagaacctcgagagggagatctactctgcaagcgtctcaagc | 866 |
| Db | 423 | AATTAGAGGACATTAATAATTGAACATGAGGGAGCGCTTTATTATTAAGAAACATTTTAAAC | 482 |
| QY | 867 | ctctgctccatctctccctccatgataagcctcttgtaagctgtgtctacacccctcaaga | 926 |
| Db | 483 | CTTAGTAGATATATCTCCAGATGCAAGCAAGCTTTCATGCTGTATATCTCTTGATATAAAA | 542 |
| QY | 927 | accggaatccatgcgcctgcctgctctcttgaacccggtgtcaggaacagtcataccatccctca | 986 |
| Db | 543 | ATAAAATCCACACATTTGCCCCGTTATTGACCCCTTACAGTGGGAATGCACATTTTATATCTTA | 602 |
| QY | 987 | cacacaagccctgcgtacaagtctctctgcaactcttggtctccctgcctgcgcgcgcctctct | 1044 |
| Db | 603 | CCCAACAAAACATCTCAAGTTCTCTCAGCTTTTATGTCTGATATGATGCCAAAGAGCGCTT | 662 |
| QY | 1047 | tcctctacgcgaactatccaaagatttgggcatcggcacatctccgaagactctgctgtgtgc | 1108 |
| Db | 663 | TCATGAAAGCGAAGACTGTGATGAGCTTGTGAATAGCAAGCTAACCAACATTTGGCTTCATAC | 722 |
| QY | 1107 | tggagacagacacccatccctgacgtgcacctggagaactcttgtagaacggcggtgtgtctgaac | 1166 |
| Db | 723 | ATCCACAGACCTCCCATCATCAAAAGCCTTCAACCATATTTTGTGAAGAAGCAACATATACGCTC | 782 |
| QY | 1167 | tgcctgtgtgtcaagaagtgtgtgtcagtgctgtggcctctacatcccgctctgtgtgtacatc | 1226 |
| Db | 783 | TGCGCTGTGGATGTAGTACGAGAAAGTTGTGATATTATTCACAAATTTGATGTATTA | 842 |
| QY | 1227 | acctgtgctcccgcaaacactcaacacacccctggaacatgagtgtggaagaagccctgaagc | 1286 |
| Db | 843 | ATCTCTCTGCTGGAATAACATCAATAAACCTTAGATATCCAGCTGACCCAGAGCGCTTTCAGC | 902 |
| QY | 1287 | agagagacatgatctctcgagggagtgctctctctgcgaagcccaagagaagccttgagggaag | 1346 |
| Db | 903 | ACCGTTACAGAGTATTTTGAAGGTGTTGTAATGTGAATTAAGCTGTGAATTTTCTTCAACCA | 962 |

| | | | |
|----|------|--|------|
| OY | 1347 | lgatcgacagatgtgtcgcggagacgaagtcacacagctgtgtgtatgtgacgaagaccaccg | 1406 |
| | | | |
| Db | 963 | ttctgtgacacatagttaagacacttcacgtccctgtgtgtgtgtaaatgaacgacata | 1022 |
| OY | 1407 | atctctctgggcgtgtgtctccctctccgaacatccctctcaagcacttgtgtctcagccctgtgt | 1466 |
| Dh | 1023 | gtattgttgctgattatttttccctctctgcgacattctgtcaaacctctgattctctacaccagcag | 1082 |
| OY | 1467 | g 1467 | |
| Db | 1083 | G 1083 | |

```

1      RESULT 4
2      : Sequence 14, Application US/08232463
3      : Patent No. 5670367
4      : GENERAL INFORMATION:
5      : APPLICANT: DORNER, F.
6      : APPLICANT: SCHEIFLINGER, F.
7      : APPLICANT: FALKNER, F. G.
8      : TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
9      : NUMBER OF SEQUENCES: 52
10     : CORRESPONDENCE ADDRESS:
11     : ADDRESSEE: Foley & Lardner
12     : STREET: 1800 Diagonal Road, Suite 500
13     : City: Alexandria
14     : STATE: VA
15     : COUNTRY: USA
16     : ZIP: 22313-0299
17     : COMPUTER READABLE FORM:
18     : MEDIUM TYPE: Floppy disk
19     : COMPUTER: IBM PC compatible
20     : OPERATING SYSTEM: PC-DOS/MS-DOS
21     : SOFTWARE: PatentIn Release #1.0, Version #1.25
22     : CURRENT APPLICATION DATA:
23     : APPLICATION NUMBER: US/08/232,463
24     : FILING DATE:
25     : CLASSIFICATION: 435
26     : PRIOR APPLICATION DATA:
27     : APPLICATION NUMBER: US/07/935,313
28     : FILING DATE:
29     : APPLICATION NUMBER: EP 91 114 300.6
30     : FILING DATE: 26-AUG-1991
31     : ATTORNEY/AGENT INFORMATION:
32     : NAME: BENT, Stephen A.
33     : REGISTRATION NUMBER: 29 768
34     : REFERENCE/DOCKET NUMBER: 30472/114 IMMU
35     : TELECOMMUNICATION INFORMATION:
36     : TELEPHONE: (703)836-9300
37     : TELEFAX: (703)683-4109
38     : TELEX: 899149
39     : INFORMATION FOR SEQ ID NO: 14:
40     : SEQUENCE CHARACTERISTICS:
41     : LENGTH: 7218 base pairs
42     : TYPE: nucleic acid
43     : STRANDEDNESS: single
44     : TOPOLOGY: linear
45     : IMMEDIATE SOURCE:
46     : CLONE: pTZipt-F15
47     : US-08-232-463-14

```

| | | | | |
|-----------------------|-------------------|-------------------|-----------|--------------|
| Query Match | 3.0% | Score 49; | DB 1; | Length 7218; |
| Best Local Similarity | 2.6% | Pred. NO. 0.0021; | | |
| Matches 10; | Conservative 220; | Mismatches 155; | Indels 0; | Gaps 0 |

```

0Y      864 agcctcctggcctccatctctctccctaaagatagcctgtttgaagctgtctaacacctcatca 923
      ||| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1057 AGCTTCGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1116
0Y      924 agaacggatccatcgctgcctgcctctctcttgaccoggtgtcaggcaacgtaactccacatcc 983

```

| | | | | |
|----------------|------|---|-----------|------|
| D _b | 1117 | yy | : : : : : | 1176 |
| O _y | 984 | tcaacacaacagcgtctcaagtcttcgtcacatttgctccctgctgccgcctt | | 1043 |
| D _b | 1177 | yy | : : : : : | 1236 |
| O _y | 1044 | ccttcctctaccgcactatccaagalltggcatcgcacattccgaacttgctgtg | | 1103 |
| D _b | 1237 | yy | : : : : : | 1296 |
| O _y | 1104 | tgctggagacagcacccatctctactgcatcaggacattctggacgcgtgtgtctg | | 1163 |
| D _b | 1297 | yy | : : : : : | 1356 |
| O _y | 1164 | cactgcctgtgtcacaagaatlgtgctcaggctgtggcctctatcccgctllgatltga | | 1223 |
| D _b | 1357 | yy | : : : : : | 1416 |
| O _y | 1224 | ttcaacctgactgccagcaacta | 1248 | |
| D _b | 1417 | yyyyyyyyyyyyyyyyyyGTACCAC | 1441 | |

```

RESULT 5
US-09-056-105-8
; Sequence 8, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; TITLE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056.105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1022
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-8

```

| | | | | |
|-----------------------|--------------|-----------------|-----------------|--------------|
| Query Match | 2.58; | Score 41.2; | DB 4; | Length 1022; |
| Best Local Similarity | 44.78; | Pred. No. 0.12; | | |
| Matches 160; | Conservative | 0; | Mismatches 198; | Indels 0; |
| | | | | Gaps 0; |

QY 373 ctgacacatcagcccttcacgttcgagctccagacagatgatgttgagcttgcacacga 4122
 Db 209 cacttgcctggaggaacccaatctgggtctccagacccaagaagaagagccacac 358
 QY 433 gtcccaagccacaagaagcctggagtgatgtgagctagaagaccctctggaagaagccctc 4322
 Db 359 ctgcgcctgaacgacagatgcctctgttccgagaagcctcagtaacaagaattgtagatctg 418
 QY 493 cctgtgcctgtcccccgaagcccatcttcccaagcttgccttgagatgacgaactgcgaa 5525
 Db 419 tcaattctctccgcaagatcatcgaccaagagctgtgtcacaaagcagaatgtctgga 478
 QY 553 accccggcgcccaagatctacatctgcgcttcacgaggaacacactgtctacatgcatgagc 612
 Db 479 gagagtcatacaaatattacaagcgtctgttctcctgtgattcttgcacaagccctccgagtc 558
 QY 613 aactagctccaagctagtatattcttgcgaacccatgctcggagatcaagaagaagccctcttgc 672
 Db 539 ccgaagaatgatctcttgcgcatctgacgtgaaggaagctggaccccgcaagcaaacctacac 598
 QY 673 tctgtgtgccaacagctgtgtggtgcagccctctatgtgacagcaagaagcagagcttt 730


```

RESULT 14
US-09-244-796-17/C
: Sequence: 17, Application US/09244796
: Patient No. 6281344
: GENERAL INFORMATION:
: APPLICANT: Szostak, Jack W.
: APPLICANT: Roberts, Richard W.
: APPLICANT: Liu, Rihc
: TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
: FILE REFERENCE: 00786/350007
: CURRENT APPLICATION NUMBER: US/09/244,796
: EARLIER FILING DATE: 1999-02-05
: EARLIER APPLICATION NUMBER: 60/035,963
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 60/064,491
: EARLIER FILING DATE: 1997-11-06
: EARLIER APPLICATION NUMBER: 09/007,005
: EARLIER FILING DATE: 1998-01-14
: NUMBER OF SEQ. ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ. ID NO: 17
: LENGTH: 289
: TYPE: RNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Translation template
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(289)
: OTHER INFORMATION: n = A,T,C or G
: US-09-244-796-17

```

[illegible]

RESULT 15
US-09-404-650-1
; Sequence 1, Application US/09404650
; Patent No. 6309838
; GENERAL INFORMATION
; APPLICANT: Dietrich, Paul S.
; APPLICANT: Mcclivern, Joseph G.
; TITLE OF INVENTION: T-TYPE CALCIUM CHANNEL VARIANTS; COMPOSITIONS THEREOF

```

? TITLE OF INVENTION: AND USES
? FILE REFERENCE: R00438-NEG sequence listing
? CURRENT APPLICATION NUMBER: US/09/404,650
? CURRENT FILING DATE: 1999-09-23
? NUMBER OF SEQ ID NOS: 12
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 6816
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (192)..(6716)
? US-09-404-650-1

```

| | | | | |
|-----------------------|----------------|----------------|----------|-------------|
| Query Match | 2.4% | Score 38.8 | DB 4 | Length 6816 |
| Best Local Similarity | 47.5% | Pred. No. 1 | | |
| Matches 115 | Conservative 0 | Mismatches 127 | Indels 0 | Gaps 0 |

| | | | |
|----|------|---|------|
| QY | 631 | catcttcgacacalgcgtgagatcaagaagcctcttgcctcgtgtgagccaagcgtgt | 690 |
| | | | |
| Db | 4169 | caacttcgcacacccctgggcccagcctctatgcttccctcttgcctcgcgtaccaaagtgtg | 4228 |
| QY | 691 | gcggcagccctcctcatgtggacagcaagaagcagcgttctgtggagatgctgcacatcac | 750 |
| | | | |
| Db | 4229 | ttgggtgacatcattgatacatgtaactgtga-tgctgtgtctgttgagccagcagcctgtgac | 4288 |
| QY | 751 | tgaacttaccctgtgtgcatactcgctactacatgaagtcgccccctcgtgcacatctatagat | 810 |
| | | | |
| Db | 4289 | caaccacacacccctggatgctgcgtctactatcatctccttcctgcgtacatcgtcaagttcct | 4348 |
| QY | 811 | tgaacaacatagattgagaccttgagaagagatctactctgaagcctcttcaaaccttc | 870 |
| | | | |
| Db | 4349 | ttgtctcaacatgtttgtgggtgtctcgtgtgtgagagaacttccaaagtctcggcagcacca | 4408 |
| QY | 871 | gg 872 | |
| | | | |
| Db | 4409 | gg 4410 | |

Search completed: October 3, 2002, 16:22:08
Job time: 16968 sec

Fri Oct 4 09:03:06 2002

us-09-826-581-5.rni

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:48:55 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280
Perfect score: 101
Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtcacc 101

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match Length | ID | Description |
|------------|-------|-------|--------------|----|-------------|
|------------|-------|-------|--------------|----|-------------|

| | | | | | | |
|----|------|-------|--------|---|-----------|---------------------|
| 1 | 101 | 100.0 | 1647 | 6 | AX281582 | AX281582 Sequence |
| 2 | 101 | 100.0 | 2290 | 9 | HS824977 | AJ24977 Homo sapi |
| 3 | 99.4 | 98.4 | 2109 | 6 | AX09776 | AX09776 Sequence |
| 4 | 99.4 | 98.4 | 2115 | 6 | AX09802 | AX09802 Sequence |
| 5 | 99.4 | 98.4 | 2115 | 6 | AF214519 | AF214519 Homo sapi |
| 6 | 72.2 | 71.5 | 1867 | 6 | AX099774 | AX099774 Sequence |
| 7 | 72.2 | 71.5 | 1873 | 4 | AF214520 | AF214520 Sus scrofa |
| 8 | 72.2 | 71.5 | 1873 | 6 | AX099800 | AX099800 Sequence |
| 9 | 72.2 | 71.5 | 2022 | 6 | AX099804 | AX099804 Sequence |
| 10 | 71 | 70.3 | 989 | 6 | AX281579 | AX281579 Sequence |
| 11 | 71 | 70.3 | 152129 | 2 | AC027416 | AC027416 Homo sapi |
| 12 | 71 | 70.3 | 206854 | 9 | AC009974 | AC009974 Homo sapi |
| 13 | 69.4 | 68.7 | 196554 | 2 | AC073128 | AC073128 Homo sapi |
| 14 | 45.4 | 45.0 | 5886 | 4 | AF214521 | AF214521 Sus scrofa |
| 15 | 42.2 | 41.8 | 227724 | 2 | AF336381 | AF336381 Mus muscu |
| 16 | 33 | 32.7 | 135012 | 8 | AP003257 | AP003257 Oryza sat |
| 17 | 32 | 31.7 | 160133 | 2 | AC024355 | AC024355 Homo sapi |
| 18 | 32 | 31.7 | 160878 | 2 | AC016971 | AC016971 Homo sapi |
| 19 | 32 | 31.7 | 162840 | 2 | AC046203 | AC046203 Homo sapi |
| 20 | 32 | 31.7 | 191330 | 9 | AC025271 | AC025271 Homo sapi |
| 21 | 32 | 31.7 | 192175 | 2 | AL591490 | AL591490 Mus muscu |
| 22 | 31.8 | 31.5 | 8670 | 8 | CRAG7 | CRAG7 Chlamydomon |
| 23 | 31.6 | 31.3 | 124057 | 9 | AP000770 | AP000770 Homo sapi |
| 24 | 31.6 | 31.3 | 129750 | 2 | AC103965 | AC103965 Homo sapi |
| 25 | 31.6 | 31.3 | 153201 | 9 | AC006454 | AC006454 Homo sapi |
| 26 | 31.6 | 31.3 | 165655 | 2 | AC027605 | AC027605 Homo sapi |
| 27 | 31.6 | 31.3 | 166941 | 2 | AC073891 | AC073891 Homo sapi |
| 28 | 31.6 | 31.3 | 194101 | 2 | AC090910 | AC090910 Homo sapi |
| 29 | 31.6 | 31.3 | 210691 | 9 | AP001891 | AP001891 Homo sapi |
| 30 | 31.4 | 31.1 | 121851 | 2 | AC091751 | AC091751 Gallus ga |
| 31 | 31.2 | 30.9 | 126380 | 2 | AP000714 | AP000714 Homo sapi |
| 32 | 31.2 | 30.9 | 139505 | 9 | HSJ591C20 | HSJ591C20 Human DNA |
| 33 | 31.2 | 30.9 | 142746 | 2 | AC093564 | AC093564 Homo sapi |
| 34 | 31.2 | 30.9 | 157473 | 2 | AL356977 | AL356977 Homo sapi |
| 35 | 31.2 | 30.9 | 159840 | 2 | AP000869 | AP000869 Homo sapi |
| 36 | 31.2 | 30.9 | 171902 | 2 | AP000846 | AP000846 Homo sapi |
| 37 | 31.2 | 30.9 | 198410 | 2 | AP000831 | AP000831 Homo sapi |
| 38 | 31.2 | 30.9 | 199321 | 9 | AP000941 | AP000941 Homo sapi |
| 39 | 31 | 30.7 | 31549 | 9 | AC096857 | AC096857 Homo sapi |
| 40 | 31 | 30.7 | 78505 | 2 | AC006408 | AC006408 Homo sapi |
| 41 | 31 | 30.7 | 123551 | 2 | AC005809 | AC005809 Homo sapi |
| 42 | 31 | 30.7 | 191832 | 2 | AC091723 | AC091723 Sus scrofa |
| 43 | 30.8 | 30.5 | 7459 | 1 | SC8A2 | AL45327 Streptomy |
| 44 | 30.8 | 30.5 | 36583 | 3 | LMFL2464 | AL65154 Tetrahym |
| 45 | 30.8 | 30.5 | 177648 | 3 | LMFP265 | AL359716 Tetrahym |

ALIGNMENTS

| RESULT | 1 | | | | | |
|------------|---|-------------|-----|--------|-----------------|--|
| AX281582 | | | | | | |
| LOCUS | AX281582 | 1647 bp | DNA | linear | PAT 02-NOV-2001 | |
| DEFINITION | Sequence 5 from Patent WO0177305. | | | | | |
| ACCESSION | AX281582 | | | | | |
| VERSION | AX281582.1 | GI:16608833 | | | | |
| KEYWORDS | | | | | | |
| SOURCE | human. | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | | |
| TITLE | Andersson, L., Luthman, H. and Marklund, S. | | | | | |
| JOURNAL | 1 (sites) | | | | | |
| FEATURES | Variaants of the human amp-activated protein kinase gamma 3 subunit | | | | | |
| source | Patent: WO 0177305-A 5 18-OCT-2001; | | | | | |
| | Arexis AB (SE) | | | | | |
| | Location/Qualifiers | | | | | |
| | 1..1647 | | | | | |
| | /organism="Homo sapiens" | | | | | |
| | /db_xref="taxon:9606" | | | | | |
| | 20..1489 | | | | | |
| | /note="unnamed protein product" | | | | | |
| | /codon_start=1 | | | | | |

| Result No. | Score | Match Length | ID | Description |
|------------|-------|--------------|----|-------------|
|------------|-------|--------------|----|-------------|

/protein_id="CAD10589.1"
/db_xref="GI:16608834"
/translation="MEPGLHRLRRTPSSLSGSEHOEMSFLEOENSSMPSPAVTS
SSERIRGKRRRAKALRWTRROKSVGEDEPPGGEGPRSPRAESNGLEATFPKTPPLAQA
DPAVGTPTPTGMDCLPSDCTASAGSTDDVELATEPRATEAMCELEGLERRPALC
LSPPAPPKLGMDELKPKGAQITMRPMOHTCTDMAATSKLVIFDTMLEIKKAFPA
LVANGVAPALWDSKOSFVGMILITDFTLVLRHYRSPVOIYEIENKHETWRELY
LQGCFKPLVLSIPNDLSLEAVYTLIKNRILHRLPYLDVSGNVHLIHLKRLKFLHIF
GSLPRPSFLYRTIODLGICTFRDLAVLETAPILTFALDIFVDRVSAALPVNVECGOV
VGLYSRFDVYHLAAQOTYNIHLDMSVGEALRORTLCLEGVLSQCPHESLGEVIDRIARE
OVHRLVLDVETOHLLGVSLSDILOALVLSFAGIDALGA"

BASE COUNT 346 a 502 c 462 g 337 t
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggccaaagccttgatgagtcagcaagcagtcgltgaggaagggagccaccagctc 60
|||||
Db 180 GGGCCAAAGCCTTGATGATGACAGCAAGCTCGTGAGAGGAGGAGCCAGCAGCTC 239
|||||
OY 61 agggggaaggtcccggtccagggcagctgtgagtcacc 101
|||||
Db 240 AGGGGGAAGGTCCCGGTCCAGGCCAGCTGCTGAGTCCACC 280
|||||

RESULT 2
HSA249977 2290 bp mRNA linear PRI 07-APR-2000
LOCUS Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
DEFINITION (AMP gamma 3 gene).
ACCESSION AJ249977
VERSION AJ249977.1 GI:6688200
KEYWORDS AMP-activated protein kinase; AMP gamma 3 gene; gamma 3 subunit.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2290)
AUTHORS Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
TITLE Characterization of AMP-activated protein kinase gamma-subunit
ISOFORMS and their role in AMP binding
JOURNAL Biochem. J. 346 Pt 3, 659-669 (2000)
MEDLINE 20164049
REFERENCE 2 (bases 1 to 2290)
AUTHORS Carling,D.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
London, W12 0NN, UNITED KINGDOM
FEATURES
source location/Qualifiers
1..2290
/organism="Homo sapiens"
/db_xref="taxon:9606"
22..1500
/gene="AMP gamma 3"
22..1500
/gene="AMP gamma 3"
/function="AMP-activated protein kinase regulatory
subunit"
/codon_start=1
/evidence="experimental
/product="AMP-activated protein kinase gamma 3 subunit"
/protein_id="CAB65117.1"
/db_xref="GI:6688201"

/translation="MEPGLHRLRRTPSSLSGSEHOEMSFLEOENSSMPSPAVTS
SSERIRGKRRRAKALRWTRROKSVGEDEPPGGEGPRSPRAESNGLEATFPKTPPLAQA
DPAVGTPTPTGMDCLPSDCTASAGSTDDVELATEPRATEAMCELEGLERRPALC
LSPPAPPKLGMDELKPKGAQITMRPMOHTCTDMAATSKLVIFDTMLEIKKAFPA
LVANGVAPALWDSKOSFVGMILITDFTLVLRHYRSPVOIYEIENKHETWRELY
LQGCFKPLVLSIPNDLSLEAVYTLIKNRILHRLPYLDVSGNVHLIHLKRLKFLHIF
GSLPRPSFLYRTIODLGICTFRDLAVLETAPILTFALDIFVDRVSAALPVNVECGOV
VGLYSRFDVYHLAAQOTYNIHLDMSVGEALRORTLCLEGVLSQCPHESLGEVIDRIARE
OVHRLVLDVETOHLLGVSLSDILOALVLSFAGIDALGA"

GSLPRPSFLYRTIODLGICTFRDLAVLETAPILTFALDIFVDRVSAALPVNVECGOV
VGLYSRFDVYHLAAQOTYNIHLDMSVGEALRORTLCLEGVLSQCPHESLGEVIDRIARE
OVHRLVLDVETOHLLGVSLSDILOALVLSFAGIDPSGPEKI"

BASE COUNT 501 a 674 c 617 g 498 t
ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 2290;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggccaaagccttgatgagtcagcaagcagtcgltgaggaagggagccaccagctc 60
|||||
Db 182 GGGCCAAAGCCTTGATGATGACAGCAAGCTCGTGAGAGGAGGAGCCAGCAGCTC 241
|||||
OY 61 agggggaaggtcccggtccagggcagctgtgagtcacc 101
|||||
Db 242 AGGGGGAAGGTCCCGGTCCAGGCCAGCTGCTGAGTCCACC 282
|||||

RESULT 3
AX099776 2109 bp DNA linear PAT 02-APR-2001
LOCUS Sequence 3 from Patent WO0120003.
DEFINITION AX099776
ACCESSION AX099776
VERSION AX099776.1 GI:13538810
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2109)
AUTHORS Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
Rogel-Gallard,C., Iannucci,N., Gellin,J., le Roy,P. and
Chardon,P.

TITLE Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof
JOURNAL Patent: WO 0120003-A 3 22-MAR-2001.
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
location/Qualifiers
1..2109
/organism="Homo sapiens"
/db_xref="taxon:9606"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35799.1"
/db_xref="GI:13538811"

/translation="MRPMOHTCTDMAATSKLVIFDTMLEIKKAFPAALVANGVRAAP
LWDSKOSFVGMILITDFTLVLRHYRSPVOIYEIENKHETWRELYLQGCFKPLVLS
IPNDLSLEAVYTLIKNRILHRLPYLDVSGNVHLIHLKRLKFLHIFGSLPRPSFL
YRTIODLGICTFRDLAVLETAPILTFALDIFVDRVSAALPVNVECGVLSRFDV
HLAAQOTYNIHLDMSVGEALRORTLCLEGVLSQCPHESLGEVIDRIAREOVHRLVLDV
TOHLLGVSLSDILOALVLSFAGIDALGA"

BASE COUNT 458 a 621 c 560 g 470 t
ORIGIN

Query Match 98.4%; Score 99.4; DB 6; Length 2109;
Best Local Similarity 99.0%; Pred. No. 4.1e-16;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gggccaaagccttgatgagtcagcaagcagtcgltgaggaagggagccaccagctc 60
|||||
Db 80 GGGCCAAAGCCTTGATGATGACAGCAAGCTCGTGAGAGGAGGAGCCAGCAGCTC 139
|||||

OY 61 agggggaaggtcccggtccagggcagctgtgagtcacc 101
|||||
Db 140 AGGGGGAAGGTCCCGGTCCAGGCCAGCTGCTGAGTCCACC 180
|||||

RESULT 4

AX099802
LOCUS AX099802 2115 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 29 from Patent WO0120003.
ACCESSION AX099802
VERSION AX099802.1 GI:13538836
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2115)
AUTHORS Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A., Rogel-Galliard,C., Iannuccelli,N., Gellin,J., le Roy,P. and Chardon,P.
TITLE Variants of the gamma chain of ampk, dna sequences encoding the same, and uses thereof
JOURNAL INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ; Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE) ; Location/Qualifiers
FEATURES
source
1..2115
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1395
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35801.1"
/db_xref="GI:13538837"
/translation="MSFLEQENSSSPSPAVTSSSERIRGRRAKALRMTRKQSVEEG
EPFGGEGRSRPTAEVSTGLEATFPKTPPLAQADPAGVGPPTGMDCLPSDCTASAG
SSTDVELATERPATEAMECELEGLERPALCLSPQAFPRKFLKMDDELKRPQAIYM
RPMOEHCTDAMATSSKLVIFDTMLEIKKAFALVANGVRAPLMDSKOSFVGMITI
TDFILVHRVRSPLVQIYEIEQHKLETRREIYLOGCFKPLVSIISPNSLFEAVYTLI
KNRIHRLPVLDPVSGNVLIHTKRLKLEFLHIFGSLRPSFLYRTIODLGIGTRDL
AVVLETPAIIALDIEVDPRVSALPVNVECGGVGLTSRFVDIHLAQQOTNHLDMSV
GEALRPTCLCEGVISCPHESLGEVIDIRAREQVHRVLVDETHLLGVSLSDILQ
ALVSPADIDALGA"
BASE COUNT 460 a 622 c 562 g 471 t
ORIGIN

Query Match 98.4% Score 99.4; DB 6; Length 2115;
Best Local Similarity 99.0% Pred. No. 4.1e-16;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggccaaagccttgagatgacaaggaagtcggtgaggaagggagccaccaggtc 60
|||||
Db 86 GGGCCAAAGCCTTGAGATGACACAGCAAGTCGGTGGAGGAGGCCACCAAGTTC 145
|||||

QY 61 aaggggaagagtcctccggtccaggaagcagctgcgtgagtcacc 101
|||||
Db 146 AGCGGGAAGCTCCCGCTCAGGCCAACCTGCTGAGTCCACC 186
|||||

RESULT 5
AF214519
LOCUS AF214519 2115 bp mRNA linear PRI 03-JUN-2000
DEFINITION Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
ACCESSION AF214519
VERSION AF214519.1 GI:8215681
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2115)
AUTHORS Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M., Rogel-Galliard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H., Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P. and Andersson,L.
TITLE A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle

JOURNAL Science 288 (5469), 1248-1251 (2000)
MEDLINE 20280150
PUBMED 10818001
REFERENCE 2 (bases 1 to 2115)
AUTHORS Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Rogel-Galliard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N., Kalm,E., le Roy,P., Chardon,P. and Andersson,L.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish University of Agricultural Sciences, BMC box 597, Uppsala 751 24, Sweden
FEATURES
source
Location/Qualifiers
1..2115
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p"
/tissue_type="skeletal muscle"
1..2115
/gene="PRKAG3"
1..1395
/gene="PRKAG3"
/note="AMPK3"
/codon_start=1
/product="AMP-activated protein kinase gamma subunit"
/protein_id="AA73987.1"
/db_xref="GI:8215682"
/translation="MSFLEQENSSSPSPAVTSSSERIRGRRAKALRMTRKQSVEEG
EPFGGEGRSRPTAEVSTGLEATFPKTPPLAQADPAGVGPPTGMDCLPSDCTASAG
SSTDVELATERPATEAMECELEGLERPALCLSPQAFPRKFLKMDDELKRPQAIYM
RPMOEHCTDAMATSSKLVIFDTMLEIKKAFALVANGVRAPLMDSKOSFVGMITI
TDFILVHRVRSPLVQIYEIEQHKLETRREIYLOGCFKPLVSIISPNSLFEAVYTLI
KNRIHRLPVLDPVSGNVLIHTKRLKLEFLHIFGSLRPSFLYRTIODLGIGTRDL
AVVLETPAIIALDIEVDPRVSALPVNVECGGVGLTSRFVDIHLAQQOTNHLDMSV
GEALRPTCLCEGVISCPHESLGEVIDIRAREQVHRVLVDETHLLGVSLSDILQ
ALVSPADIDALGA"
BASE COUNT 460 a 622 c 562 g 471 t
ORIGIN

Query Match 98.4% Score 99.4; DB 9; Length 2115;
Best Local Similarity 99.0% Pred. No. 4.1e-16;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggccaaagccttgagatgacaaggaagtcggtgaggaagggagccaccaggtc 60
|||||
Db 86 GGGCCAAAGCCTTGAGATGACACAGCAAGTCGGTGGAGGAGGCCACCAAGTTC 145
|||||

QY 61 aaggggaagagtcctccggtccaggaagcagctgcgtgagtcacc 101
|||||
Db 146 AGCGGGAAGCTCCCGCTCAGGCCAACCTGCTGAGTCCACC 186
|||||

RESULT 6
AX099774
LOCUS AX099774 1867 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 1 from Patent WO0120003.
ACCESSION AX099774
VERSION AX099774.1 GI:13538808
KEYWORDS pig.
SOURCE
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 1867)
AUTHORS Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A., Rogel-Galliard,C., Iannuccelli,N., Gellin,J., le Roy,P. and Chardon,P.
TITLE Variants of the gamma chain of ampk, dna sequences encoding the same, and uses thereof
JOURNAL INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ; Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)

| FEATURES | |
|----------|------------------------|
| source | Location/Qualifiers |
| | 1. 1867 |
| | /organism="Sus scrofa" |
| | /db_xref="taxon:9823" |
| CDS | 472. . 1389 |

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 380 a | 583 c | 529 g | 375 t |
| ORIGIN | | | | |

| | | | | |
|-----------------------|-----------------|--------------------|-----------|--------------|
| Query Match | 71.5%; | Score 72.2; | DB 6; | Length 1867; |
| Best Local Similarity | 82.2%; | Pred. No. 4.6e-09; | | |
| Matches 83; | Conservative 0; | Mismatches 18; | Indels 0; | Gaps 0; |

[illegible]

| RESULT | 7 |
|------------|--|
| LOCUS | AF214520 |
| DEFINITION | AF214520 1873 bp mRNA linear MAM 03-JUN-2000 |
| | Sus scrofa AMP-activated protein kinase gamma subunit (PRKAC3) |
| | mRNA, complete cds. |

| | | |
|----------|-------------------|------------|
| VERSION | AF214520.1 | GI:8215683 |
| KEYWORDS | | |
| SOURCE | pig. | |
| ORGANISM | <i>Sus scrofa</i> | |

REFERENCE
AUTHORS
Milan, D., Jeon, J.-T., Looft, C., Amarver, V., Rohic, A., Thelander M.
1 (bases 1 to 1873)
Euryota, Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 1873)

| Author | Title |
|---|--|
| Lundström, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P. L., Chardon, P. and Andersson, L. | A mutation in PFKF3 associated with excess glycogen content in pig skeletal muscle |

MEDLINE 20280150
 PUBMED 10818001
 2 (bases 1 to 1873)
 REFERENCE
 AUTHORS
 Milian D. Toes TM. 1996b. 3

TITLE
Direct Submision
Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
University of Agricultural Sciences, BMC box 597, uppsala 751 24,
Sweden

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .1873 |

```

"organism"="Sus scrofa"
"db_xref"="taxon:9823"
"chromosome"="15"
"map"="15q"
"tissue_type"="skeletal muscle"
1. .1873
"gene"="PRKAG3"
"gene"="PRKAG3"
1. .1395
CDS

```

```

/note="AMPG3"
/codon_start=1
/product="AMP-activated protein kinase gamma subunit"
/protein_id="AAFP398.1"
/db_xref="GI:821668"

```

| BASE COUNT | 382 a | 580 c | 535 g | 376 t |
|------------|-------|-------|-------|-------|
| ORIGIN | | | | |

| | | | | |
|--------------------------|--------|--------------------|-----------|--------------|
| Query Match | 71.5% | Score 72.2; | DB 4; | Length 1873; |
| Best Local Similarity | 82.2%; | Pred. No. 4.6e-09; | | |
| Matches 83; Conservative | 0; | Mismatches 18; | Indels 0; | Gaps 0; |

[illegible]

| | | |
|------------|------------------------|-----------------|
| Accession | AX099800.1 | GI:13538834 |
| Version | Sequence 27 | |
| Definition | from Patent WO0120003. | |
| Locus | AX099800 | |
| Size | 1873 bp | DNA |
| Type | linear | PAT 02-APR-2001 |

| KEYWORDS | SOURCE | ORGANISM |
|----------|------------|------------------------------|
| plg. | Sus scrofa | Eukaryota; Metazoa; Chordata |

REFERENCE
AUTHORS
1 Mammalia: Euteria; Cetartiodactyla; Suina; Suidae; Sus.
(bases 1 to 1873)
Anderson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Rogel-Gallard, C., Iannuccelli, N., Gellin, J., Le Roy, P. and

TITLE Variants of the gamma chain of ampx, dna sequences encoding the same, and uses thereof
PATENT: WO 0120003-A 27 22-MAR-2001;
JOURNAL INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FRI):

| FEATURES | Location/Qualifiers | Source |
|----------|---------------------|--------|
| | 1. .1873 | |

```
/organism="Sus scrofa"  
/db_xref="taxon:9823"  
1. .1395
```

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35800.1"
/db_xref="GI:13538835"
/translation="MSRLDQGSBSPEPSRANTTSSESHQDGNKASRMTROEDVEEG
GPGCPREDGQSPVAEESQDEATFFPKATPLQAPAPLAEVNPPTREDDILPSDCAASAS
DSNTHLDLGIETSSASADGDELGVKRPAPCPSPVLLPRLGMDELDLQAPDAVYM
HFQHEQETIDANATSKSLVIFEDTMELETAKFAFLVANGVAAALMSKQSFQVGLTI
TDLVLVHRTYRSPLOVIELEHRIETWRIYVLDGCFPLVSIIPNDISI.FAAVVALI
KNNIHLPLADPVSAGVALIHTKRLKLFLEIPICTLLPRSPSYRTIOLDGIGTRDL
AAVLEAPLITLALDIFVDRVSALPVVNETQVGLDSRPDIYVLAQOQTYNNHLMGN
GEMLRORTLCGEVLVSCQPHETLGEVIDRIYREQVHRLVLYDEDTQHLGVTSIDTIO
ALVSPAGIDALGA"

```

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 382 a | 580 c | 535 g | 376 t |
| ORIGIN | | | | |

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135376 bases at least Q40
Consensus quality: 143264 bases at least Q30
Consensus quality: 146503 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 149029; sum-of-contigs
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1006 1105: gap of 100 bp in length
1106 2402: contig of 1297 bp in length
2403 2502: gap of 100 bp
2503 3823: contig of 1321 bp in length
3824 3923: gap of 100 bp
3924 5020: contig of 1097 bp in length
5021 5120: gap of 100 bp
5121 6161: contig of 1041 bp in length
6162 6261: gap of 100 bp
6262 7547: contig of 1286 bp in length
7548 7647: gap of 100 bp
7648 9983: contig of 2336 bp in length
9984 10083: gap of 100 bp
10084 12556: contig of 2473 bp in length
12557 12656: gap of 100 bp
12657 15043: contig of 2387 bp in length
15044 15143: gap of 100 bp
15144 17123: contig of 1980 bp in length
17124 17223: gap of 100 bp
17224 19466: contig of 2243 bp in length
19467 19566: gap of 100 bp
19567 21928: contig of 2362 bp in length
21929 22028: gap of 100 bp
22029 24319: contig of 2291 bp in length
24320 24419: gap of 100 bp
24420 27059: contig of 2640 bp in length
27060 27159: gap of 100 bp
27160 30170: contig of 3011 bp in length
30171 30270: gap of 100 bp
30271 33968: contig of 3698 bp in length
33969 34068: gap of 100 bp
34069 38179: contig of 4111 bp in length
38180 38279: gap of 100 bp
38280 42366: contig of 4087 bp in length
42367 42466: gap of 100 bp
42467 46365: contig of 3899 bp in length
46366 46465: gap of 100 bp
46466 51285: contig of 4820 bp in length
51286 51385: gap of 100 bp
51386 55871: contig of 4486 bp in length
55872 55971: gap of 100 bp
55972 60595: contig of 4524 bp in length
60596 60695: gap of 100 bp
60696 66595: contig of 5900 bp in length
66596 66695: gap of 100 bp
66696 73218: contig of 6523 bp in length
73219 73318: gap of 100 bp
73319 77115: contig of 3797 bp in length
77116 77215: gap of 100 bp
77216 85022: contig of 7807 bp in length
85023 85122: gap of 100 bp
85123 93314: contig of 8192 bp in length
93315 93414: gap of 100 bp
93415 101193: contig of 7779 bp in length

* 101194 101293: gap of 100 bp
* 101294 113090: contig of 11797 bp in length
* 113091 113190: gap of 100 bp
* 113191 123496: contig of 10306 bp in length
* 123497 123596: gap of 100 bp
* 123597 137837: contig of 14241 bp in length
* 137838 137937: gap of 100 bp
* 137938 152129: contig of 14192 bp in length.

FEATURES

source

misc_feature
1. 1005
/note="assembly-fragment"
misc_feature
1106..2402
/note="assembly-fragment"
misc_feature
2503..3823
/note="assembly-fragment"
misc_feature
3924..5020
/note="assembly-fragment"
misc_feature
5121..6161
/note="assembly-fragment"
misc_feature
6262..7547
/note="assembly-fragment"
misc_feature
7648..9983
/note="assembly-fragment"
misc_feature
10084..12556
/note="assembly-fragment"
misc_feature
12657..15043
/note="assembly-fragment"
misc_feature
15144..17123
/note="assembly-fragment"
misc_feature
17224..19466
/note="assembly-fragment"
misc_feature
clone_end:SP6
vector_side:left"
misc_feature
19567..21928
/note="assembly-fragment"
misc_feature
22029..24319
/note="assembly-fragment"
misc_feature
24420..27059
/note="assembly-fragment"
misc_feature
27160..30170
/note="assembly-fragment"
misc_feature
30271..33968
/note="assembly-fragment"
misc_feature
34069..38179
/note="assembly-fragment"
misc_feature
38280..42366
/note="assembly-fragment"
misc_feature
42467..46365
/note="assembly-fragment"
misc_feature
46466..51285
/note="assembly-fragment"
misc_feature
51386..55871
/note="assembly-fragment"
misc_feature
55972..60595
/note="assembly-fragment"
misc_feature
60696..66595
/note="assembly-fragment"
misc_feature
66696..73218
/note="assembly-fragment"
misc_feature
73319..77115
/note="assembly-fragment"
misc_feature
77216..85022
/note="assembly-fragment"
misc_feature
85123..93314
/note="assembly-fragment"
misc_feature
93415..101193
/note="assembly-fragment"
misc_feature
101294..113090

/note="assembly_fragment"

Query Match 70.3%: Score 71; DB 2; Length 152129;
 Best Local Similarity 100.0%: Pred. No. 3.8e-09;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gggccaaagccttgatgacaaagcagaagtcggtgaggaaggagccacagtc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2945 GGCCCAAGCCTTGAGATGGACAAAGCAGAGTCGGTGGAGGAAGGAGCCAGCAGTC 2886
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Oy 61 aagggaggaagtc 71
 ||||||||||||
 Db 2885 ACCGCGAAGCT 2875

RESULT 12
 AC009974/c 206854 bp DNA linear PRI 09-JAN-2002
 LOCUS AC009974 Homo sapiens BAC clone RP11-459119 from 2, complete sequence.
 DEFINITION AC009974
 AC009974
 VERSION AC009974.9 GI:16799058
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 206854)
 Sulton, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99063792

REFERENCE 2 (bases 1 to 206854)
 Hariri, A. and Cotton, M.
 The sequence of Homo sapiens BAC clone RP11-459119
 Unpublished (2001)

REFERENCE 3 (bases 1 to 206854)
 Waterston, R.H.
 Direct Submission
 Submitted (08-SEP-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 4 (bases 1 to 206854)
 Waterston, R.H.
 Direct Submission
 Submitted (08-NOV-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 5 (bases 1 to 206854)
 Waterston, R.H.
 Direct Submission
 Submitted (03-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 6 (bases 1 to 206854)
 Waterston, R.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 8, 2001 this sequence version replaced gi:13431203.

COMMENT ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0459119

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPc1-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
 Tateo, M., Catanese, J.J., and de Jong, P.J. (1996) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu)
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-1077K22; the clone
 sequenced to the right is RP11-64705. Actual start of this clone
 is at base position 1 of RP11-459119; actual end is at base
 position 206854 of RP11-459119.

Data from AC079810 and AC073128 was used to finish this clone,
 AC009974. Polymorphisms have been identified between AC073128 and
 AC009974. A single plasmid region exists between 38812-38903. An
 unresolved tandem in the HERV SVA exists between 184390-185163.
 PCR suggests that approximately 1700 bps are missing.

FEATURES

source

1..206854

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-459119"

/clone_11b="RPC1-11"

1..37

/note="match to EST BE314060 (NID:g9134719)"

1..37

/note="match to EST BI059713 (NID:g14467240)"

1..37

/note="match to EST BF183086 (NID:g11061273)"

1..37

/note="match to EST AL567345 (NID:g12920610)"

1..37

/note="match to EST AW880850 (NID:g8042860)"

1..37

/note="match to EST BF304755 (NID:g11251653)"

1..37

/note="similar to Homo sapiens EST B1114348
(NID:g14565249)"

1..37

/note="match to EST BG477625 (NID:g13409904)"

1..37

/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

3..37

/note="match to EST BE908408 (NID:g10402954)"

4..37

/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

164..662

/note="match to EST C05773 (NID:g1502549)"

220..221

/note="match to EST BG470047 (NID:g13402322)"

281..344

/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

281..344

misc_feature

```

misc_feature /note="match to EST BE908408 (NID:g10402954)"
281..344 /note="match to EST BF304755 (NID:g11251653)"
misc_feature 281..344
misc_feature /note="match to EST BF183086 (NID:g11061273)"
281..344 /note="similar to Homo sapiens EST B1114348
(NID:g14565249)"
misc_feature 281..344
misc_feature /note="match to EST BG477625 (NID:g13409904)"
281..344 /note="match to EST B1059713 (NID:g14467240)"
misc_feature 281..344
misc_feature /note="match to EST AL567345 (NID:g12920610)"
281..344 /note="match to EST BE314060 (NID:g91347719)"
281..344 /note="match to EST A1670836 (NID:g4850567) wa04g10.x1"
281..344 /note="match to EST AW880850 (NID:g8042860)"
294..344 /note="match to EST BG470047 (NID:g13402322)"
complement(57..558)
misc_feature /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
588..929 /note="match to EST T64073 (NID:g667938) yc05d12.r1"
594..763 /note="match to EST BF304755 (NID:g11251653)"
594..763 /note="match to EST BF304755 (NID:g13402322)"
594..763 /note="match to EST BC470047 (NID:g13402322)"
594..763 /note="match to EST AW880850 (NID:g8042860)"
594..731 /note="match to EST BE314060 (NID:g91347719)"
594..764 /note="match to EST B1059713 (NID:g14467240)"
594..763 /note="match to EST BC477625 (NID:g13409904)"
594..742 /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
594..763 /note="match to EST BE908408 (NID:g10402954)"
594..763 /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
594..763 /note="match to EST A1670836 (NID:g4850567) wa04g10.x1"
594..763 /note="match to EST BF183086 (NID:g11061273)"
594..763 /note="similar to Homo sapiens EST B1114348
(NID:g14565249)"
599..763 /note="match to EST AL567345 (NID:g12920610)"
617..1084 /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
622..763 /note="match to EST A1860958 (NID:g5514574) w156f05.x1"
622..763 /note="similar to Mus musculus EST A1196847 (NID:g3749453)
u167e01.x1"
684..763 /note="match to EST BG992568 (NID:g14396638)"
962..1084 /note="match to EST A1656812 (NID:g4740791) tc54b06.x1"
967..1084 /note="match to EST BE908408 (NID:g10402954)"
967..1085 /note="match to EST BF304755 (NID:g11251653)"
967..1091 /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
967..1071
misc_feature
Query Match 70.3% Score 71 DB 9 Length 206854

```

```

Best Local Similarity 100.0% Pred. No. 3.6e-09;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gggccaaagcccttgagatgacaaagccagaagtgcgttgaggaaaggggcccacccagtc 60
|||||
Db 168897 GGCCAAAGCCTTGACATGACAAAGCAGCAAGCTGCTGGAGGAAGGAGCCACGACGTC 168838

Oy 61 aggggggaaggt 71
|||||
Db 168837 AGGGGCAAGGT 168827

RESULT 13
AC073128/c
AC073128/c
LOCUS
DEFINITION
AC073128.3 GI:13027579
AC073128
VERSION
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 196554)
The sequence of Homo sapiens clone
Waterston,R.H.
Unpublished
2 (bases 1 to 196554)
Waterston,R.H.
Direct Submission
Submitted (08-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 21, 2001 this sequence version replaced g1:8469048.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0647005
----- Summary Statistics -----
Sequencing vector: M13; 98%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 98% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990119
Consensus quality: 187795 bases at least Q40
Consensus quality: 190513 bases at least Q30
Consensus quality: 192099 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 194954; sum-of-contigs
Quality coverage: 5.58 in Q20 bases; agarose-fp
Quality coverage: 5.67 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1157: contig of 1157 bp in length
* 1158 1257: gap of unknown length
* 1258 3600: contig of 2343 bp in length
* 3601 3700: gap of unknown length
* 3701 5103: contig of 1403 bp in length
* 5104 5203: gap of unknown length
* 5204 8524: contig of 3321 bp in length
* 8525 8624: gap of unknown length
* 8625 11856: contig of 3232 bp in length

```

```
* 11957 11956: gap of unknown length
* 11957 15783: contig of 3827 bp in length
* 15784 15883: gap of unknown length
* 15884 21906: contig of 6023 bp in length
* 21907 22006: gap of unknown length
* 22007 28887: contig of 6881 bp in length
* 28888 35255: gap of unknown length
* 35256 35355: gap of unknown length
* 35356 44642: contig of 9287 bp in length
* 44643 44743: gap of unknown length
* 44743 58375: contig of 13533 bp in length
* 58376 58376: gap of unknown length
* 58376 73816: contig of 15441 bp in length
* 73817 73916: gap of unknown length
* 73917 92140: contig of 18224 bp in length
* 92141 92240: gap of unknown length
* 92241 113337: contig of 21097 bp in length
* 113338 130325: contig of 16888 bp in length
* 130326 130425: gap of unknown length
* 130426 149287: contig of 18862 bp in length
* 149288 149387: gap of unknown length
* 149388 196554: contig of 47167 bp in length.
FEATURES
source
Location/Qualifiers
1..196554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-64705"
1..1157
/feature="assembly_name:Contig17"
1258..3600
/feature="assembly_name:Contig18"
3701..5103
/feature="assembly_name:Contig19
vector_side:right"
5204..8524
/feature="assembly_name:Contig20"
8625..11856
/feature="assembly_name:Contig21"
11957..15783
/feature="assembly_name:Contig22"
15884..21906
/feature="assembly_name:Contig23"
22007..28887
/feature="assembly_name:Contig24"
28988..35255
/feature="assembly_name:Contig25"
35356..44642
/feature="assembly_name:Contig26"
44743..58375
/feature="assembly_name:Contig27"
58376..73816
/feature="assembly_name:Contig28"
73917..92140
/feature="assembly_name:Contig29"
92241..113337
/feature="assembly_name:Contig30"
113438..130325
/feature="assembly_name:Contig31"
130426..149287
/feature="assembly_name:Contig32"
149388..196554
/feature="assembly_name:Contig33
clone_end:SP6
vector_side:right"
BASE COUNT 52296 a 46993 c 45889 g 49770 t 1606 others
ORIGIN
Query Match 68.7% Score 69.4 DB 2 Length 196554
```

```
Best Local Similarity 98.6% Pred. No. 9.4e-09;
Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 gggccaaagccttgagatgacagaagtcggtggaaggaagggagccaccagtc 60
|||||
Db 63318 GGGCCAAAGCCTTGAGATGACAGAGCAGAACTCGCTGAGAGGAGGAGCAGACAGTC 63259
|||||
Oy 61 agggggaaggt 71
|||||
Db 63258 AGGGGGAAGGT 63248

RESULT 14
AF214521 5888 bp DNA linear MAM 03-JUN-2000
LOCUS
DEFINITION
ACCESSION
AF214521
VERSION
AF214521.1 GI:8215685
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 5888)
Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
Rogel-Gallard,C., Paul,S., Iannuccelli,N., Raek,L., Bonne,H.,
Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.,
and Anderson,L.
A mutation in PRKG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)
20280150
PUBMED
10818001
2 (bases 1 to 5888)
Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
Rogel-Gallard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
Direct Submission
Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
Sweden
FEATURES
source
Location/Qualifiers
1..5888
/organism="Sus scrofa"
/db_xref="taxon:9823"
/chromosome="15"
/map="15q"
join(<1..154,515..918,1809..1890,2349..2407,2509..2554,
2771..2825,3027..3153,3286..3451,4578..4615,4791..4937,
5294..5410)
/feature="PRKG3"
/product="AMPK gamma subunit"
<1..5410
/feature="PRKG3"
join(1..154,515..918,1809..1890,2349..2407,2509..2554,
2771..2825,3027..3153,3286..3451,4578..4615,4791..4937,
5294..5410)
/feature="PRKG3"
/codon_start=1
/product="AMPK gamma subunit"
/protein_id="AAF33989.1"
/db_xref="GI:8215686"
/translation="MSFLDEGSRSPRAVTTSSERSHDGDNKASRMTROEDVEG
GPGRPRGPGSRPVAESTGQATFPKATPLAQAPLAELVDNPTERRDILPSDCAASAS
DSMTDHLDLGIIEFSASASGDELGLVEKPAQSPSEVLLPRGMDDELQPCGAQVYM
HFMQERTCYDAMATSSKLVIEDTMLKKAFALVANGVRAAPLMSKOSQPGMGLTI
TDFILVLRHYRSPLOVYIEIEHEKLEIEMKEIYLQCCFPVVISNSDLSFAYVALI
KNRIHRLPVLDIPVSGAVLHILTKKRLKELHITGTLIPRPSFLIRTIQDLGIGTRFDL
AVYLEAPRLITLADIDFVDRVSAALPVYNETGVVGLSRFDVTHLAQQTYNRLDMNV
GEALRQRTLCLEGVLSQPHETLGEVIDRIVRQVRLVLDVETQHLGLCVLSLDTLQ
ALVLSAPGIDALCA"
BASE COUNT 1274 a 1634 c 1638 g 1339 t 3 others
ORIGIN
```

PN WO200177305-A2.
XX 18-OCT-2001.
XX 06-APR-2001; 2001WO-SE00765.
XX 07-APR-2000; 2000US-195665P.
XX (AREX-) AREXIS AB.
XX Andersson L, Luthman H, Marklund S;
PI WPI: 2001-657170/75.
DR P-PSDB: Q0B47679.
XX
XX New variants of human AMP-activated protein kinase gamma3 subunit
PT associated with a metabolic disease e.g. diabetes or obesity and method
PT for determining a risk estimate of diseases in subject by detecting the
PT variant -
XX
XX Disclosure: Fig 5; 25pp; English.
XX
XX This sequence represents the full length cDNA encoding the human
CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC a risk estimate of a metabolic disease, such as diabetes or obesity,
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon
CC 3 variation may be a substitution of a G for a C at nucleotide 320,
CC resulting in the amino acid substitution P71A; in exon 4 variation may
CC be a substitution of a T for a C at nucleotide 550; and in exon 10
CC variation may be a substitution of a T for a C at nucleotide 1037,
CC resulting in the amino acid substitution R340W. There may also be
CC nucleotide variation in intron 6. The numbering of these
CC variations is based on the full length cDNA as given, rather than on
CC position 1 of the open reading frame.
XX
XX Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;
SQ

Query Match 100.0%; Score 101; DB 22; Length 1647;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gggccaagccttgagatgagcaagaagtcggtgaggaaggagccacaggtc 60
DB 180 gggccaagccttgagatgagcaagaagtcggtgaggaaggagccacaggtc 239
OY 61 agggggaaggtcccggtccagcgagcagctgctgagtcacc 101
DB 240 agggggaaggtcccggtccagcgagcagctgctgagtcacc 280

RESULT 2
AAD03296
ID AAD03296 standard; DNA; 2109 BP.
XX
XX AAD03296:
XX 13-JUN-2001 (first entry)
XX
XX Human AMK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
DE
XX Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..471
FT /*tag= a
FT CDS 472..1389

FT /*tag= b
FT /product= "Human Prkag3 protein"
FT 3'UTR 1390..2109
FT /*tag= c
XX
XX WO200120003-A2.
XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-EP09896.
XX
XX 10-SEP-1999; 99EP-0402236.
XX 18-MAY-2000; 2000EP-0401388.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFT C.
XX (KALM/) KALM E.
XX
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Galliard C,
PI Iannuccielli N, Gellin J, Le Roy P, Chardon P;
PI WPI: 2001-244810/25.
DR P-PSDB: AAE00221.
XX
XX New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
XX Claim 12: Fig 2; 71pp; English.
XX
XX The present sequence is a cDNA encoding human adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC useful as therapeutic for treating carbohydrate metabolism disorders such
CC as diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder.
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
XX Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;
SQ

Query Match 98.4%; Score 99.4; DB 22; Length 2109;
Best Local Similarity 99.0%; Pred. No. 4.1e-20;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gggccaagccttgagatgagcaagaagtcggtgaggaaggagccacaggtc 60
DB 80 gggccaagccttgagatgagcaagaagtcggtgaggaaggagccacaggtc 139
OY 61 agggggaaggtcccggtccagcgagcagctgctgagtcacc 101
DB 140 agggggaaggtcccggtccagcgagcagctgctgagtcacc 180

RESULT 3
AAD03320
ID AAD03320 standard; cDNA; 2115 BP.
XX

AC AAD0320;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
 XX
 KW Human: gamma subunit; adenosine monophosphate-activated kinase; AMPK;
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
 KW cystathione beta synthase; CBS; cardiant; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1395
 FT /tag= a
 FT /product= "Human complete Prkag3 protein"
 XX
 PN W0200120003-A2.
 XX
 PD 22-MAR-2001.
 XX
 PP 11-SEP-2000; 2000WO-EP09896.
 XX
 PR 10-SEP-1999; 99EP-0402236.
 PR 18-MAY-2000; 2000EP-0401388.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA (ANDE/) ANDERSSON L.
 PA (LOOF/) LOOFT C.
 PA (KALM/) KALM E.
 XX
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
 PI Iannuccielli N, Gellin J, Le Roy P, Chardon P;
 DR WPI: 2001-244810/25.
 DR P-PSDB: AAE00223.
 XX
 DT New variants of the gamma subunit of vertebrate adenosine
 DT monophosphate-activated kinase for diagnosis or treatment of disorders
 DT associated with energy metabolism such as diabetes, obesity, and
 DT myopathy -
 XX
 PS Claim 12: Page 65-68; 71pp: English.
 XX
 CC The present sequence is a cDNA encoding human adenosine monophosphate
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
 CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
 CC useful as therapeutic for treating carbohydrate metabolism disorders such
 CC as diabetes, obesity, and disorders associated with muscle metabolism
 CC such as myopathy and cardiovascular diseases, to modulate AMPK
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
 CC and its functionally altered mutants are useful for the diagnostic
 CC evaluation, genetic testing and prognosis of a metabolic disorder,
 CC preferably a carbohydrate metabolism disorder. Primers that can detect
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting
 CC from the expression of a functionally altered allele of PRKAG3.
 CC Transgenic animal and host cell transformed with PRKAG3 or a
 CC heterotimeric AMPK consisting of PRKAG3 or its mutant, are useful for
 CC screening compounds able to modulate AMPK activity. Nucleic acid
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain
 CC of PRKAG3 and is useful in gene therapy.
 XX
 SO Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;

Query Match 98.4%; Score 99.4; DB 22; Length 2115;
 Best Local Similarity 99.0%; Pred. No. 4; le-20;
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 gggccaaagccttgatgacgaagcagaagtcgtgaggaagggaaccaccagtc 60
 DB 86 gggccaaagccttgatgacgaagcagaagtcgtgaggaagggaaccaccagtc 145
 OY 61 agggggaaggtcccggtccaggcagctgctgctcacc 101
 DB 146 agggggaaggtcccggtccaggcagctgctgctcacc 186

RESULT 4
 AAD03295
 ID AAD03295 standard; CDNA: 1867 BP.
 XX
 AC AAD03295;
 XX
 DT 13-JUN-2001 (first entry)
 DT
 DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
 XX
 KW Pig: gamma subunit; adenosine monophosphate-activated kinase; AMPK;
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
 KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
 KW chromosome 15; ss.
 XX
 OS Sus scrofa.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..471
 FT /tag= a
 FT 472..1389
 FT CDS /tag= b
 FT /product= "Sus scrofa PRKAG3 protein"
 FT 3'UTR 1390..1867
 FT /tag= c
 XX
 PN W0200120003-A2.
 XX
 PD 22-MAR-2001.
 XX
 PP 11-SEP-2000; 2000WO-EP09896.
 XX
 PR 10-SEP-1999; 99EP-0402236.
 PR 18-MAY-2000; 2000EP-0401388.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA (ANDE/) ANDERSSON L.
 PA (LOOF/) LOOFT C.
 PA (KALM/) KALM E.
 XX
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
 PI Iannuccielli N, Gellin J, Le Roy P, Chardon P;
 DR WPI: 2001-244810/25.
 DR P-PSDB: AAE00220.
 XX
 DT New variants of the gamma subunit of vertebrate adenosine
 DT monophosphate-activated kinase for diagnosis or treatment of disorders
 DT associated with energy metabolism such as diabetes, obesity, and
 DT myopathy -
 XX
 PS Claim 12: Fig 2: 71pp: English.
 XX
 CC The present sequence is a cDNA encoding pig adenosine monophosphate
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
 CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
 CC Mutation in Prkag3 results in an altered regulation of carbohydrate
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
 CC therapeutic for treating carbohydrate metabolism disorders such as
 CC diabetes, obesity, and disorders associated with muscle metabolism
 CC such as myopathy and cardiovascular diseases, to modulate AMPK
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
 CC and its functionally altered mutants are useful for the diagnostic

XX 10-JAN-2000 (first entry)
 DT XX
 DE Human METH2 related EST X16619.
 XX
 KW Human: METH1; METH2: anti-angiogenic; metalloprotease thrombospondin;
 KW cancer: diagnosis; hyperproliferative disorder; autoimmune disease;
 KW angiogenesis inhibitor; abnormal wound healing; inflammation;
 KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
 KW diabetic retinopathy; macula degeneration; haemangioma; detection;
 KW arterial-venous malformation; immune deficiency; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO937660-A1.
 PD 29-JUL-1999.
 XX
 PF 22-JAN-1999; 99WO-US01313.
 XX
 PR 23-JAN-1998; 98US-0072298.
 PK 28-AUG-1998; 98US-0098539.
 XX
 PA (IRUE/) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 XX
 PI Irue-la-Arispe L, Hastings GA, Ruben SM;
 XX
 DR WPI: 1999-590684/50.
 XX
 PT New isolated metalloprotease thrombospondin polypeptides, useful for
 PT treating hyperproliferative disorders, cancers or autoimmune disorders
 PT -
 PS Disclosure: Page 431-437; 457pp; English.
 XX
 CC AA232000 and AA232001 encode, and AAY49501 and AAY49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC endometrial bleeding disorders, diabetic retinopathy, some forms of
 CC macula degeneration, haemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilisation (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. AA232002 to AA232080, and AAY49503 to
 CC AAY49511 represent sequences given in the exemplification of the present
 CC invention.
 XX
 SO Sequence 8670 BP; 1624 A; 2421 C; 3266 G; 1359 T; 0 other;

Query Match 31.5%; Score 31.8; DB 20; Length 8670;
 Best Local Similarity 64.0%; Pred. No. 3.8;
 Matches 48; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 14 gagatgacaagcagaagtcggtgaggaagggagccaccaggtcaggggaggaagtc 73
 DB 6582 GAGGTGAGCGTGGAGCGGCGTGGAGAGCGTGGAGCGGTATCATGAACAGCC 6523
 OY 74 ccggcgcagggcagc 88
 DB 6522 CCCGCCAACCCCGC 6508

RESULT 9

AAC90316/c
 ID AAC90316 standard; DNA; 8670 BP.
 XX
 AC AAC90316;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE X16619 cDNA clone.
 XX
 KW METH; metalloprotease; thrombospondin; angiogenesis inhibition;
 KW cancer therapy; benign tumour; ocular angiogenic disease;
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
 KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
 KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
 KW coronary collateral; cerebral collateral; arteriovenous malformation;
 KW ischaemic limb angiogenesis; Osler-Weber syndrome; wound granulation;
 KW plaque neovascularisation; telangiectasia; haemophilic joint; EST;
 KW angiodiroma; fibromuscular dysplasia; expressed sequence tag;
 KW Crohn's disease; atherosclerosis; birth control; ss.
 XX
 OS Unidentified.
 XX
 PN WO200071577-A1.
 PD 30-NOV-2000.
 XX
 PF 25-MAY-2000; 2000WO-US14462.
 XX
 PR 25-MAY-1999; 99US-0318208.
 PR 20-JUL-1999; 99US-0144882.
 PR 10-AUG-1999; 99US-0147823.
 PR 13-AUG-1999; 99US-0373658.
 PR 22-DEC-1999; 99US-0171503.
 PR 22-FEB-2000; 2000US-0183792.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (SMIK-) SMITHKLINE BEECHAM CORP.
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 PA (IRUE/) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 PA (JONAV/) JONAK 2 L.
 PA (TRUL/) TRULLI S H.
 PA (FORN/) FORMALD J A.
 PA (TERR/) TERRETT J A.
 XX
 PI Irue-la-Arispe L, Hastings GA, Ruben SM, Jonak 2L, Trulli SH;
 PI Fornwald JA, Terrett JA;
 XX
 WPI: 2001-025136/03.
 XX
 PT METH1 and METH2 polynucleotides and encoded polypeptides, used to
 PT inhibit angiogenesis in the treatment of disorders such as cancer,
 PT rheumatoid arthritis and psoriasis -
 PT
 XX
 XX Claim 14: Pages 731-737; 768pp; English.
 XX
 CC The present invention relates to human METH1 and METH2, (ME for
 CC metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
 CC The present sequence is an expressed sequence tag (EST) for METH. METH
 CC can be used for inhibiting angiogenesis in an individual, and for
 CC treating cancer, benign tumours, an ocular angiogenic disease,
 CC rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
 CC vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
 CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
 CC coronary collateral, cerebral collateral, arteriovenous malformations,
 CC ischaemic limb angiogenesis, Osler-Weber syndrome, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma,
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or
 CC atherosclerosis. METH can also be used in birth control. METH can also
 CC be used in diagnostic methods for the prognosis of cancer.
 XX
 SO Sequence 8670 BP; 1624 A; 2421 C; 3266 G; 1359 T; 0 other;

sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from Wind.

and to produce other types of data and products dependent on own seed

amino acid sequences. AAS64197-AAS94564 represent novel human

PT primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX
XX
PS Claim 1: SEQ ID 3676; 2537pp + CD ROM: English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 886 BP: 169 A; 259 C; 213 G; 241 T; 4 other;

Query Match 29.1%; Score 29.4; DB 22; Length 886;
Best Local Similarity 58.6%; Pred. No. 14;
Matches 51: Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 2 ggcacaaagccttgatgacagaagtcgltggaaggaagggaccacagatca 61
DB 100 GCCTAAGCACTCCGTAGGCGCCAGCGCGAGCGCTCGAGAGCAAGCGCC 41
OY 62 gggggaaggtcccggtccagggcagc 88
DB 40 AGAGAGAAAGTCCCGGGGCGCAGCTTC 14

RESULT 14
AAH14551/C
ID AAH14551 standard; cDNA: 1457 BP.

XX
AC AAH14551;

XX
DT 26-JUN-2001 (first entry)

XX
DE Human cDNA sequence SEQ ID NO:12117.

XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS
XX Homo sapiens.

XX
PN EP1074617-A2.

XX
PD 07-FEB-2001.

XX
PE 28-JUL-2000; 2000EP-0116126.

XX
PR 29-JUL-1999; 99JP-0248036.

XX
PR 27-AUG-1999; 99JP-0300253.

XX
PR 11-JAN-2000; 2000JP-0118776.

XX
PR 02-MAY-2000; 2000JP-0183767.

XX
PR 09-JUN-2000; 2000JP-0241899.

XX
PA (HELI-) HELIX RES INST.

XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.

PT primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 8: SEQ ID 12117; 2537pp + CD ROM: English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 1457 BP: 291 A; 442 C; 365 G; 359 T; 0 other;

Query Match 29.1%; Score 29.4; DB 22; Length 1457;
Best Local Similarity 58.6%; Pred. No. 15;
Matches 51: Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 2 ggcacaaagccttgatgacagaagtcgltggaaggaagggaccacagatca 61
DB 100 GCCTAAGCACTCCGTAGGCGCCAGCGCGAGCGCTCGAGAGCAAGCGCC 41
OY 62 gggggaaggtcccggtccagggcagc 88
DB 40 AGAGAGAAAGTCCCGGGGCGCAGCTTC 14

RESULT 15
AAV59752/C
ID AAV59752 standard; DNA: 1482 BP.

XX
AC AAV59752;

XX
DT 19-JAN-1999 (first entry)

XX
DE Human secreted protein gene 96 clone HAQBK61.

XX
XX

XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;

XX
KW Immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX
OS Homo sapiens.

```

PN      MO9839448-AZ.
XX      11-SEP-1998.
XX      06-MAR-1998: 98WO-US04493.
XX      02-OCT-1997: 97US-0061060.
PR      07-MAR-1997: 97US-0038621.
XX      07-MAR-1997: 97US-0040161.
PR      07-MAR-1997: 97US-0040162.
PR      07-MAR-1997: 97US-0040163.
PR      07-MAR-1997: 97US-0040333.
PR      07-MAR-1997: 97US-0040334.
PR      07-MAR-1997: 97US-0040336.
PR      07-MAR-1997: 97US-0040337.
PR      11-APR-1997: 97US-0043311.
PR      11-APR-1997: 97US-0043312.
PR      11-APR-1997: 97US-0043313.
PR      11-APR-1997: 97US-0043569.
PR      11-APR-1997: 97US-0043576.
PR      11-APR-1997: 97US-0043578.
PR      11-APR-1997: 97US-0043580.
PR      11-APR-1997: 97US-0043659.
PR      11-APR-1997: 97US-0043670.
PR      11-APR-1997: 97US-0043671.
PR      11-APR-1997: 97US-0043672.
PR      11-APR-1997: 97US-0043674.
PR      23-MAY-1997: 97US-0047492.
PR      23-MAY-1997: 97US-0047500.
PR      23-MAY-1997: 97US-0047501.
PR      23-MAY-1997: 97US-0047502.
PR      23-MAY-1997: 97US-0047503.
PR      23-MAY-1997: 97US-0047581.
PR      23-MAY-1997: 97US-0047582.
PR      23-MAY-1997: 97US-0047583.
PR      23-MAY-1997: 97US-0047584.
PR      23-MAY-1997: 97US-0047585.
PR      23-MAY-1997: 97US-0047586.
PR      23-MAY-1997: 97US-0047587.
PR      23-MAY-1997: 97US-0047588.
PR      23-MAY-1997: 97US-0047589.
PR      23-MAY-1997: 97US-0047590.
PR      23-MAY-1997: 97US-0047592.
PR      23-MAY-1997: 97US-0047593.
PR      23-MAY-1997: 97US-0047594.
PR      23-MAY-1997: 97US-0047595.
PR      23-MAY-1997: 97US-0047596.
PR      23-MAY-1997: 97US-0047597.
PR      23-MAY-1997: 97US-0047598.
PR      23-MAY-1997: 97US-0047599.
PR      23-MAY-1997: 97US-0047600.
PR      23-MAY-1997: 97US-0047601.
PR      23-MAY-1997: 97US-0047612.
PR      23-MAY-1997: 97US-0047613.
PR      23-MAY-1997: 97US-0047614.
PR      23-MAY-1997: 97US-0047615.
PR      23-MAY-1997: 97US-0047617.
PR      23-MAY-1997: 97US-0047618.
PR      23-MAY-1997: 97US-0047632.
PR      23-MAY-1997: 97US-0047633.
PR      06-JUN-1997: 97US-0048964.
PR      06-JUN-1997: 97US-0048974.
PR      13-JUN-1997: 97US-0049610.
PR      16-JUL-1997: 97US-0051926.
PR      18-AUG-1997: 97US-0052874.
PR      22-AUG-1997: 97US-0055724.
PR      22-AUG-1997: 97US-0056630.
PR      22-AUG-1997: 97US-0056631.
PR      22-AUG-1997: 97US-0056632.
PR      22-AUG-1997: 97US-0056636.
PR      22-AUG-1997: 97US-0056637.

```

```

PR      22-AUG-1997: 97US-0056662.
PR      22-AUG-1997: 97US-0056664.
PR      22-AUG-1997: 97US-0056845.
PR      22-AUG-1997: 97US-0056862.
PR      22-AUG-1997: 97US-0056864.
PR      22-AUG-1997: 97US-0056872.
PR      22-AUG-1997: 97US-0056874.
PR      22-AUG-1997: 97US-0056875.
PR      22-AUG-1997: 97US-0056876.
PR      22-AUG-1997: 97US-0056877.
PR      22-AUG-1997: 97US-0056878.
PR      22-AUG-1997: 97US-0056879.
PR      22-AUG-1997: 97US-0056880.
PR      22-AUG-1997: 97US-0056881.
PR      22-AUG-1997: 97US-0056882.
PR      22-AUG-1997: 97US-0056884.
PR      22-AUG-1997: 97US-0056886.
PR      22-AUG-1997: 97US-0056887.
PR      22-AUG-1997: 97US-0056888.
PR      22-AUG-1997: 97US-0056889.
PR      22-AUG-1997: 97US-0056903.
PR      22-AUG-1997: 97US-0056908.
PR      22-AUG-1997: 97US-0056909.
PR      22-AUG-1997: 97US-0056910.
PR      22-AUG-1997: 97US-0056911.
PR      05-SEP-1997: 97US-0057650.
PR      05-SEP-1997: 97US-0057651.
PR      05-SEP-1997: 97US-0057659.
PR      12-SEP-1997: 97US-0057761.
PR      12-SEP-1997: 97US-0058785.

PA      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI      Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI      Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI      Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
XX      MPI; 1998-506364/43.
DR      P-PSDB; AAW74967.
XX
XX      New isolated human genes and the secreted polypeptide(s) they encode
PT      - useful for diagnosis and treatment of e.g. cancers, neurological
PT      disorders, immune diseases, inflammation or blood disorders
XX
XX      Claim 1: Page 478-479; 721pp: English.
PS
PS      This sequence represents a nucleic acid molecule designated Gene 96 from
XX      CC the human CDNA clone HA0BK61 (deposited as clone ATCC 97897 and ATCC
XX      CC 209043) which encodes a secreted human protein. The gene can be used to
XX      CC generate fusion proteins by linking to the gene to a human
XX      CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
XX      CC the fused protein as compared to the human protein only.
XX      CC The invention relates to 186 novel genes and their fragments (nucleic
XX      CC acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
XX      CC which are useful for preventing, treating or ameliorating medical
XX      CC conditions e.g. by protein or gene therapy. Also, pathological
XX      CC conditions can be diagnosed by determining the amount of the new
XX      CC polypeptides in a sample or by determining the presence of mutations in
XX      CC the new polynucleotides. Specific uses are described for each of the 186
XX      CC polynucleotides, based on which tissues they are most highly expressed in
XX      CC (see AAV59511 for described uses).
SQ      Sequence 1482 BP; 308 A; 444 C; 368 G; 362 T; 0 other:

```

```

Query Match          29 1%: Score 29.4; DB 19: Length 1482;
Best Local Similarity 58.6%: Pred. No. 15;
Matches 51: Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Qy      2 ggcacaaagccttgagatgcacaaagcagaagtcggtgaggaaggagacacacagtcga 61

```

Fri Oct 4 09:03:08 2002

us-09-826-581-5_copy_180_280.rng

Page 11

Pb 90 GGCCTAAGCACTCCCTTATGGGGCCAGGGGCCAACGGCGGAGGCTCTGGAGGAGACAGAGGCC 31
 QY 62 qggqgaaggtccccgggtcccaagcagc 88
 Db 30 AGACAGAGAGTCCCGGGGGCCACCTTGC 4

Search completed: October 3, 2002, 16:30:43
Job time: 14313 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:01 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280

Perfect score: 101

Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtcacac 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues 27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_liv:*
15: em_gss_pln:*
16: em_gss_vrc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 67.4 | 66.7 | 399 | 9 | AM356079 |
| 2 | 67.4 | 66.7 | 422 | 9 | AM427435 |
| 3 | 67.4 | 66.7 | 621 | 9 | BB628877 |
| 4 | 67.4 | 66.7 | 653 | 9 | BB629521 |
| 5 | 65.8 | 65.1 | 685 | 9 | BB630381 |
| 6 | 54.6 | 54.1 | 548 | 10 | BI775360 |
| 7 | 53.8 | 53.3 | 516 | 9 | AI664508 |
| 8 | 52.6 | 52.1 | 444 | 10 | BF890374 |
| 9 | 35.6 | 35.2 | 922 | 10 | BI754053 |
| 10 | 32.6 | 32.3 | 406 | 9 | BB759831 |
| 11 | 32.6 | 32.3 | 457 | 9 | BB780864 |
| 12 | 32.6 | 32.3 | 736 | 10 | BI765906 |
| 13 | 32.4 | 32.1 | 151 | 10 | BG244994 |
| 14 | 32 | 31.7 | 454 | 9 | BB860266 |
| 15 | 32 | 31.7 | 479 | 9 | BF549656 |
| 16 | 31.2 | 30.9 | 896 | 10 | BF526844 |
| 17 | 31 | 30.7 | 327 | 9 | BB753946 |

| | | | | | |
|----|------|------|------|----|----------|
| 18 | 31 | 30.7 | 1074 | 10 | BG966577 |
| 19 | 30.8 | 30.5 | 365 | 9 | AA787203 |
| 20 | 30.8 | 30.5 | 491 | 10 | BF651383 |
| 21 | 30.8 | 30.5 | 722 | 12 | A2188883 |
| 22 | 30.6 | 30.3 | 512 | 9 | AI939950 |
| 23 | 30.6 | 30.3 | 1195 | 10 | BE299786 |
| 24 | 30.4 | 30.1 | 370 | 10 | BG086356 |
| 25 | 30.4 | 30.1 | 443 | 9 | BB845787 |
| 26 | 30.4 | 30.1 | 514 | 10 | BI181375 |
| 27 | 30.2 | 29.9 | 167 | 10 | BF393187 |
| 28 | 30.2 | 29.9 | 211 | 9 | BB438733 |
| 29 | 30.2 | 29.9 | 428 | 10 | BF394635 |
| 30 | 30.2 | 29.9 | 545 | 9 | AA413109 |
| 31 | 30.2 | 29.9 | 825 | 10 | BI910794 |
| 32 | 30.2 | 29.9 | 1079 | 12 | CNS05J0V |
| 33 | 30 | 29.7 | 250 | 10 | BE426030 |
| 34 | 30 | 29.7 | 409 | 10 | BE426831 |
| 35 | 30 | 29.7 | 429 | 10 | BE489634 |
| 36 | 30 | 29.7 | 1253 | 11 | AK015101 |
| 37 | 29.8 | 29.5 | 335 | 10 | 225209 |
| 38 | 29.8 | 29.5 | 385 | 12 | A0852870 |
| 39 | 29.8 | 29.5 | 381 | 12 | BH414079 |
| 40 | 29.8 | 29.5 | 662 | 12 | CNS048SD |
| 41 | 29.8 | 29.5 | 740 | 10 | BF135550 |
| 42 | 29.8 | 29.5 | 1101 | 12 | CNS00004 |
| 43 | 29.6 | 29.3 | 285 | 9 | BB039663 |
| 44 | 29.6 | 29.3 | 375 | 9 | AA298840 |
| 45 | 29.6 | 29.3 | 431 | 9 | BB822055 |

ALIGNMENTS

RESULT 1
LOCUS AM356079 399 bp mRNA linear EST 25-APR-2001
DEFINITION 38073 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.
ACCESSION AM356079
VERSION AM356079.1 GI:6860085
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 399)

REFERENCE Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Perten,G., Holt,L., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013

TITLE JOURNAL MEDLINE COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68913-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 17 row: P column: 9
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..399
/organism="Bos taurus"
/db_xref="taxon:9913"

FEATURES

source

```

/c1ome.lib="MARC 280v"
/c1ssue_type="pooled"
/lab_host="PH10B"
/notes="Vector: pCMV SPORT6; Site-1: XbaI; Site-2: XhoI
Library made from pooled tissue from testis, thymus,
semilandoncus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT 100 a 126 c 117 g 56 t
ORIGIN

```

| | | | | | | | |
|-----------------------|-------|--------------|---------|------------|----|--------|-----|
| Query Match | 66.7% | Score | 67.4 | DB | 9 | Length | 359 |
| Best Local Similarity | 79.2% | Pred. No. | 1.8e-07 | | | | |
| Matches | 80 | Conservative | 0 | Mismatches | 21 | Indels | 0 |
| | | | | | | Gaps | 0 |

[illegible]

| | |
|------------|---|
| RESULT | 2 |
| AM427435 | |
| LOCUS | 422 bp mRNA linear |
| DEFINITION | 63185 MARC 3BOV Bos taurus cDNA 5' , mRNA sequence. |
| ACCESSION | AM427435 |
| VERSION | AM427435.1 GI:955382 |
| KEYWORDS | EST . |
| SOURCE | COW . |

REFERENCE
1 (bases 1 to 422)
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 422)

TITLE
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013

JOURNAL
JOURNAL
MEDLINE
COMMENT

CONTACT: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -m1nscore 20
and -minmatch 12 options.

FEATURES
SOURCE

| BASE COUNT | ORIGIN |
|------------|------------------|
| 102 a | 136 c 125 g 59 t |

| | | | | |
|--------------------------|--------|--------------------|-----------|-------------|
| Query Match | 66.78; | Score 67.4; | DB 9; | Length 422; |
| Best Local Similarity | 79.28; | Pred. No. 1.8e-07; | | |
| Matches 80; Conservative | 0; | Mismatches 21; | Indels 0; | Gaps 0; |

Oy 1 ggGCCAAAGccttggaatgacaaagcgagaatgcgttggaggaaagggagccaccagtc 60
|| ||| | | ||||| | | | | | ||||| |
Db 170 GGACCAAGGTCCTCAAGATGAAAAGCCAGAGATGTACAAGGAAGGGAGACTGCCACGCC 229

| | |
|----|---|
| Qy | 61 agggggaagctcccggtccaggcagctgtgattccaac 101
 |
| Db | 230 TGAAGGAGATCCCACTCAGGCAGACTGTCTGAGTCCACC 270 |

| RESULT | 3 |
|---|---|
| BB628877 | |
| LOCUS | BB628877 |
| DEFINITION | 621 bp mRNA linear EST 31-AUG-2001 |
| BB628877 | BB628877 |
| BB628877 | BB628877 |
| musculus cdna clone 9630039422 5', mRNA sequence. | musculus cdna clone 9630039422 5', mRNA sequence. |
| BB628877 | BB628877 |

| | | |
|----------|----------------------|-------------|
| VERSION | BB628877.1 | GI:15399483 |
| KEYWORDS | EST. | |
| SOURCE | house mouse. | |
| ORGANISM | mouse (Mus musculus) | |

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 621)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

TITLE Matsuura, T., Miyazaki, A., Nomura, K., Ohno, M., Sasak
JOURNAL Okiaki, Y., Okido, T., Saito, R., Sakai, K., Sano, H., Sasak
COMMENT , D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

URL: <http://genome.gsc.riken.go.jp/Carcinip/cp/Shibata,Y.,Hayatsu,N.,Sugahara,Y.,Shibata,K.,Itoh,M.,Konno,H.,Okazaki,Y.,Moramatsu,M.andHayashizaki,Y.>
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagii, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T., Matsura S., Kawai, J., Okazaki, Y., Moramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000).

Kondo, H., Fukunishi, T., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shimagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, T., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y. Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Funct. Genomics* 2 pre, L72-L76 (2001).

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .621 |

| Query Match | 54.1% | Score 54.6 | DB 10 | Length 548 |
|-----------------------|-------|--|-------|---------------|
| Best Local Similarity | 77.6% | Pred. No. 0.00033 | | |
| Matches | 66 | Conservative | 0 | Mismatches 19 |
| | | | | Indels 0 |
| | | | | Gaps 0 |
| QY | 17 | algcagagcagaagtcgclqgaagaaaggaagcaccaggtcagggggaaggtcccg | 76 | |
| | | | | |
| | | | | |
| Db | 33 | ATGCAGAAAGCCAGGAGGATGTAACGAAAGGAGCTGCCAGGCTGGAGGAGGTCCCA | 92 | |
| | | | | |
| QY | 77 | gllccaggcagctgctgaagtcacc | 101 | |
| | | | | |
| Db | 93 | TTTCAGAGCAGCTCTGAGTCCACC | 117 | |

| | |
|------------|---|
| RESULT | 7 |
| A1664508 | |
| LOCUS | A1664508 |
| DEFINITION | A166A508 516 bp mRNA linear EST 10-MAY-1999 |
| ACCESSION | U825005.y1 Sugano mouse embryo mewa Mus musculus cDNA clone |
| VERSION | A1664508 IMAGE:1970001 5', mRNA sequence. |
| KEYWORDS | A1664508.1 GI:4768091 |
| SOURCE | EST. |
| ORGANISM | house mouse. |
| | Mus musculus |

| REFERENCE | AUTHORS | |
|--------------------|---|---|
| 1 (bases 1 to 516) | Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., | Maniatis, T.; Eubleria; Rodentia; Sciurognathi; Muridae; Murinae; Mus |

TITLE: The WashU-NCI Mouse EST Project 1999
JOURNAL: Unpublished (1999)

Contact: Maria M.Maslin-MCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LINT; contact the
IMAG Consortium (InfoImage.llnl.gov) for further information.
MGI:986741

seq primer: custom primer used
High quality sequence stop: 484.

| source | location/qualifiers |
|---------|---------------------|
| 1. .516 | |

```

/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970001"
/clone_1lb="Sugano mouse embryo meka"
/dev_stage="embryo, 14 dpc"
/lab_host="D1110B"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TTTGGGCTACTGTG), digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCGCTCTAAAGCTGCG and 3' end primer CGAGCTGACGCTCGACGCA."

```

Query Match

53.38; Score 53.8; DB 9; Length 516;

| | | | | | |
|---------|-----------------------|--|-------|-------------|-----------------------|
| | Best Local Similarity | 77.2% | Pred. | No. 0.00051 | |
| Matches | 78: | Conservative | 0: | Mismatches | 22: Indels 1; Gaps 1: |
| OY | 1 | gggccaaagcccttgcgatacgacaacaggcaagaatcgtgtgaagaaagggaaccaccagtgc | 60 | | |
| | | | | | |
| Dd | 408 | GAGTAAAGGCTTCACGAATGGACGACAAGCAGAGCCCGGTAGAGCAAMACCAGAACCATCAGTT | 467 | | |
| OY | 61 | aaggggaaggtcccgcgtgccaggccagactgtctaatccacc | 101 | | |
| | | | | | |
| Dd | 468 | TGGGAGAGGGTGCCA-GTCCACAGACCAGCGTGCATTCACC | 507 | | |

[illegible]

REFERENCE
1 (bases 1 to 44)
Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T.,

| TITLE | JOURNAL | DATE |
|--|------------------------------------|----------|
| Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle | Genome Res. 11 (4), 626-630 (2001) | 21/06/03 |

COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Smith@ars.usda.gov
Virology Pass Sequencing. Bases called and ALT-trimmed with phred
v090904.e. Vector identified by cross_match with the -mnscore 18
and minmatch 12 options.

RESULTS

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .444 |

```

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1fb="MARC 3BOV"
/tissue_type="pooled"
/ab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."

```

Query Match 52.18; Score 52.6; DB 10; Length 444;

Best Local Similarity 81.38; Pred. No. 0.00099;
Matches 61; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

27 cagaagtcgctgcaaggaaggagccaccaggtcagggggaagtcgccggtccagcgcca 86

87 gctgctgagtcacc 101

| Db | 61 | GCTGCTGATGTCAC | 75 |
|-----------------------|--|------------------------------------|---|
| RESULT | 9 | B1754053 | 922 bp mRNA linear EST 25-SEP-2001 |
| LOCUS | B1754053/c | 603027607f1 NIH_MGC_114 | Homo sapiens CDNA clone IMAGE:5198016 5', |
| DEFINITION | | mRNA sequence. | |
| ACCESSION | B1754053 | | |
| VERSION | B1754053.1 | GI:15745631 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| TITLE | 1 (bases 1 to 922) | | |
| JOURNAL | NIH-MGC http://mgc.ncl.nih.gov/ . | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11495 row: n column: 01
High quality sequence stop: 830m.
Location/Qualifiers
1..922
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5198016"
/clone_1id="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library." | | |
| BASE COUNT | 216 a 268 c 284 g 154 t | | |
| ORIGIN | | | |
| Query Match | 35.2%; Score 35.6; DB 10; Length 922; | | |
| Best Local Similarity | 62.2%; Pred. No. 23; | | |
| Matches | 56; Conservative 0; Mismatches 34; Indels 0; Gaps 0; | | |
| Qy | 4 ccaagaccttaagtcagcgaagtcgctggggaagggaagcaccacagtcagg 63 | | |
| Db | 915 CGACAGGTTTCCCGCCGAGTGGTTCGATTTCCACAGTGGAGCTGAGCCACAGGCCCT 856 | | |
| Qy | 64 gggaaggtcccggtccagagccagctgctg 93 | | |
| Db | 855 GGGAGTGGGCCCGTTCCGTCGCGACCTGCTG 826 | | |
| RESULT | 10 | | |
| LOCUS | B8759831/c | 406 bp mRNA linear EST 17-OCT-2001 | |
| DEFINITION | B8759831 RIKEN full-length enriched, melanocyte Mus musculus CDNA | | |
| ACCESSION | B8759831 | | |
| VERSION | B8759831.1 | GI:16206159 | |
| KEYWORDS | EST. | | |
| SOURCE | house mouse. | | |
| ORGANISM | Mus musculus | | |

| | | | | |
|-----------------------|---|--|--|--|
| REFERENCE | Mammalia: Eutelia: Chordata: Craniata: Vertebrata: Euteleostomi: Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutelia: Rodentia: Sciurognathi: Muridae: Murinae: Mus 1 (bases 1 to 406) | | | |
| AUTHORS | Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishi,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Salto,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,T., Takehashi,F., Takaku-Ahira,S., Tanaka,T., Tomaru,A., Toyota,T., Watabiki,A., Yasunishi,A., Yamashita,M., Hayashizaki,Y., RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001) | | | |
| TITLE | Unpublished (2001) | | | |
| JOURNAL | Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/ | | | |
| COMMENT | Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagik,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyama,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues. | | | |
| FEATURES | Location/Qualifiers
1..406
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G27012L23"
/clone_11b="RIKEN full-length enriched, melanocyte"
/cell_type="melanocyte" | | | |
| BASE COUNT | 77 a 130 c 98 g 101 t | | | |
| ORIGIN | | | | |
| Query Match | 32.3%; Score 32.6; DB 9; Length 406; | | | |
| Best Local Similarity | 58.9%; Pred. No.1,1e+02; | | | |
| Matches | 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0; | | | |
| QY | 7 aagcttgaagatgcacaagcagaagtcgtgtaggaaggaaggagccaccaggtcagggg 66

Db 257 AAGCCCTGATGTTCCCAAGCAGAAAGCGAGGAGGAGAAAGCAGAGCCAGAGTGGAGG 198 | | | |
| QY | 67 aaggtcccggtccagcagcagctgtgtgtcacc 101

Db 197 CAGGAGCCCGCAGCAGGCTCTCTATCAAGTTCAAGC 163 | | | |
| RESULT 11 | 457 bp mRNA linear EST 15-NOV-2001 | | | |
| LOCUS | BB780864 RIKEN full-length enriched, colon RCB-0349 C1e-H3 cDNA Mus | | | |
| DEFINITION | musculus cDNA clone G43008010.3, mRNA sequence. | | | |

ACCESSION B0780864
VERSION B0780864.1 GI:16941564
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus. 1 (bases 1 to 457)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Nomasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanabe,A., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y. RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
TITLE Unpublished (2001)
JOURNAL Contact: Yoshihide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
carninci.p., shibata.y., hayatsu.n., sugahara.y., shibata.k., itoh.m., kono.h., okazaki.y., muramatsu.m., and hayashizaki.y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
waji.k., fujiwaka.s., inoue.k., togawa.y., izawa.m., ohara.e., watabiki.m., yoneda.y., ishikawa.t., ozawa.k., tanaka.t., matsura.s., kawai.j., yoneda.y., ishikawa.t., muramatsu.m., inoue.y., kira.a., and hayashizaki.y.
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.
FEATURES
source location/Qualifiers
1..457
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G430068010"
/clone.lib="RIKEN full-length enriched, colon RCB-0549
c1e-H3 cDNA"
/tissue_type="colon"
/cell_line="RCB-0549 C1e-H3"
BASE COUNT 93 a 145 c 106 g 113 t
ORIGIN
Query Match 32.3%: Score 32.6; DB 9; Length 457;
Best Local Similarity 58.9%: Pred. No. 1,1e+02;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 7 aagccttgatgacgaagcgaagtcggtgaggaaggaagcaccaggtcagggg 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 313 AACGCTGAGTTCCCAACAGACGACGAGGAGGAGGAGGAGGAGGAGGAGG 254
QY 67 aagctcccgagccagggcagagctgcaggtcagtcacac 101
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 253 AAGGACCCGAGACGAGTCTGCTATCAACTTCACG 219

RESULT 12
LOCUS B1765906/c 736 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046160F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186367 5', mRNA sequence.
ACCESSION B1765906
VERSION B1765906.1 GI:15757484
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiinae: Homo. 1 (bases 1 to 736)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L14M11465 row: h column: 16
High quality sequence stop: 736.
FEATURES
source location/Qualifiers
1..736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5186367"
/clone.lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH-MGC Library."
BASE COUNT 138 a 213 c 194 g 191 t
ORIGIN
Query Match 32.3%: Score 32.6; DB 10; Length 736;
Best Local Similarity 60.9%: Pred. No. 1.2e+02;
Matches 53; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 2 ggcacaaagctgagtgacgaagcgaagtcggtgaggaaggaagcaccaggtca 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 113 GGCCTAAGACTCCCTTAGGGCCAGCGGCAACGCGGAGCGCTCGACCGACGAGGCC 54
QY 62 gggggaaggtcccggtcagtcagccagc 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53 AGACAGAAAGTCCCGGGGCGGCGACGCTCGACCGACGAGGCC 27
RESULT 13
LOCUS BG244994 151 bp mRNA linear EST 13-FEB-2001
DEFINITION 602358260F1 NCI-CCAP_Mam1 Mus musculus cDNA clone IMAGE:4486719 5', mRNA sequence.
ACCESSION BG244994
VERSION BG244994.1 GI:12754809
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|--------------------|---|--|--------------------|-----------------------------------|
| 1 (bases 1 to 151) | Mammalian; eutheria; | modestia; | sciuromathi; | Muridae; Murinae; Mus |
| NIH-MGC | http://mgc.nci.nih.gov/ . | National Institutes of Health, Mammalian Gene Collection (MGC) | Unpublished (1999) | Contact: Robert Strausberg, Ph.D. |

Tissue Procurement: Gilbert Smith, Ph.D.,
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10330 row: h column: 16
High quality sequence stop: 151.

| FEATURES | location/Qualifiers |
|----------|---------------------|
| SOURCE | 1. .151 |

| | | | | |
|------------|---------------------------------------|------|------|------|
| BASE COUNT | 21 a | 38 c | 70 g | 22 t |
| ORIGIN | Providing samples: Gilbert Smith, NIH | | | |

| | | | | |
|-----------------------|-----------------|----------------|-----------|-------------|
| Query Match | 32.1% | Score 32.4; | DB 10; | Length 151; |
| Best Local Similarity | 68.2%; | Pred. No. 97; | | |
| Matches 45; | Conservative 0; | Mismatches 21; | Indels 0; | Gaps 0; |

OY 14 gagatgacaaaggcagaagtcggltgaggaagygagccaccagtlcagggggaagtcc 73
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 GAGTGGACATCGGCACACGCCGCTGAGAGGTGACGATCACCCCGGAAGGGGGGACGACG 125

| | | | |
|----|-----|--------|-----|
| QY | 74 | ccggtc | 79 |
| | | | |
| Db | 126 | TGGGTC | 131 |

| RESULT | LOCUS | DEFINITION | Accession |
|--------|----------|---|-----------|
| 14 | BB860266 | 454 bp mRNA linear EST 26-NOV-2001 | |
| | BB860266 | BB860266 RIKEN full-length enriched, pooled cell lines Mus musculus | |
| | | cDNA clone G430012004 5' mRNA sequence. | |

| REFERENCE | AUTHORS | TITLE | JOURNAL |
|--------------------|---|--|--------------------|
| 1 (bases 1 to 444) | Akimura, T., Atakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishikawa, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Ahikira, S., Tanabe, Y., Tomaru, A., Toyu, T., Watanabe, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | Riken Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001) | Unpublished (2001) |
| Contact: | Yoshihide Hayashizaki | | |

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shuho-cho, Tsukuba, Ibaraki 305-3858, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichipillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000).

Komoto, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .454 |

| BASE COUNT | 121 a | 123 c | 133 g | 77 t |
|------------|-------|-------|-------|------|
| ORIGIN | | | | |

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 31.7% | Score 32; | DB 9; | Length 454; |
| Best Local Similarity | 60.2%; | Pred. No. 1.6e+02; | | |
| Matches 53; | Conservative 0; | Mismatches 35; | Indels 0; | Gaps 0; |

6 aaagccttgatgacaaagcaagaatcggtggaaggggagaccacccagtccaagg 65
| | | | | | | | | | | | | | | |
Db 357 ACACCATAGAGATGTC AAGCCAAAGATCAGCACGAAGAAGGGGATCCCCCTTGACCACGA 416

Db 417 GAGCTGATCTTGCAGGCAAGCAGCTG 444

| RESULT | 15 |
|------------|---|
| BP549656 | |
| LOCUS | |
| DEFINITION | 479 bp mRNA linear EST 12-DEC-2000 |
| ACCESSION | U1-R-C2-ng-f-01-0-01 .r1 U1-R-C2 Rattus norvegicus cDNA clone |
| | U1-R-C2-ng-f-01-0-01 5', mRNA sequence. |
| | BP549656 |

```

: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: FILING DATE: 14-FEB-1997
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,385
: FILING DATE: 03-AUG-1994
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09905
: FILING DATE: 03-AUG-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/776,859
: FILING DATE: 03-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: BROOK, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: MSP97-01A2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 617-861-9540
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 268 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: US-08-800-840-1
:
: Query Match      28.3%; Score 28.6; DB 3; Length 268;
: Best Local Similarity 55.6%; Pred. No. 2.6;
: Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
:
: QY 1 gggccaaagccttgagatgacaaagcagagtcgltgaggaaggagccaccagtc 60
:    | |||| | || | || | || | || | || | || | || | || | || | || |
: DB 50 GAGCCACACCTTAGGTTGGCCATCTACTCCAGAGCAGGAGGAGGAGCCAGGCGC 109
:
: QY 61 agggggaaggtcccggtccaggtccaggtcgatgcca 99
:    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 110 TGGGCATTAAGTCAAGGCAGAGCCATCTATTGCTTACA 148
:
: RESULT 3
: US-08-776-859-1
: Sequence 1, Application US/08776859
: Patent No. 6090592
: GENERAL INFORMATION:
: APPLICANT: Mosaic Technologies, Inc.
: APPLICANT: Adams, Christopher P.
: APPLICANT: Kiron, Stephen J.
: TITLE OF INVENTION: Method and Apparatus for Performing
: FILE REFERENCE: Amplification of Nucleic Acids on Supports
: CURRENT APPLICATION NUMBER: US/08/776,859
: FILING DATE: 1997-05-29
: EARLIER APPLICATION NUMBER: 08/285,385
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 268
: TYPE: DNA
: ORGANISM: Homo sapiens
:
: US-08-776-859-1
```

```

: Query Match      28.3%; Score 28.6; DB 3; Length 268;
: Best Local Similarity 55.6%; Pred. No. 2.6;
: Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
:
: QY 1 gggccaaagccttgagatgacaaagcagagtcgltgaggaaggagccaccagtc 60
:    | |||| | || | || | || | || | || | || | || | || | || | || |
: DB 50 GAGCCACACCTTAGGTTGGCCATCTACTCCAGAGCAGGAGGAGGAGCCAGGCGC 109
:
: QY 61 agggggaaggtcccggtccaggtccaggtcgatgcca 99
:    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
: DB 110 TGGGCATTAAGTCAAGGCAGAGCCATCTATTGCTTACA 148
:
: RESULT 4
: PCT-US95-09905-1
: Sequence 1, Application PC/TUS9509905
: GENERAL INFORMATION:
: APPLICANT:
: APPLICANT:
: APPLICANT:
: TITLE OF INVENTION: METHOD AND APPARATUS FOR PERFORMING
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
: STREET: 600 ATLANTIC AVENUE
: CITY: BOSTON
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09905
: FILING DATE: FILED HEREWITH
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/285,385
: FILING DATE: 03-AUGUST-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: JANIUK, ANTHONY J.
: REGISTRATION NUMBER: 29,809
: REFERENCE/DOCKET NUMBER: B0833/7001MO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 268 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
:
: PCT-US95-09905-1
:
: Query Match      28.3%; Score 28.6; DB 5; Length 268;
: Best Local Similarity 55.6%; Pred. No. 2.6;
: Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
:
: QY 1 gggccaaagccttgagatgacaaagcagagtcgltgaggaaggagccaccagtc 60
:    | |||| | || | || | || | || | || | || | || | || | || | || |
: DB 50 GAGCCACACCTTAGGTTGGCCATCTACTCCAGAGCAGGAGGAGGAGCCAGGCGC 109
:
: QY 61 agggggaaggtcccggtccaggtccaggtcgatgcca 99
```


TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 11
PCT-US96-09430-7

Query Match 28.3% Score 28.6; DB 5; Length 6060;
Best Local Similarity 55.6%; Pred. No. 4.7;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Y 1 gggccaaagccttgagatgagacaaagcagaagtcgtgaggaaggagccacacagtc 60
Db 2041 GAGCCACACCCCTAGGCTTGCCCAATCTACTCCAGGAGCAGGAGGAGGAGCCAGGCC 2100
Y 61 agggggaaggtcccggtccagggcagctgctgagtcga 99
Db 2101 TGGGCTATAAGTCAAGGAGGAGCCATCTATTCCTTACA 2139

RESULT 11
US-08-289-653-2/c
Sequence 2, Application US/08289653
Patent No. 5543322
GENERAL INFORMATION:
APPLICANT: Kazuaki KITANO et al.
TITLE OF INVENTION: DNA AND ITS USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,653
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/887,284
FILING DATE: May 22, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Fusarium sp.

STRAIN: S-19-5 (IFO 8884)
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-289-653-2

Query Match 26.5% Score 26.8; DB 1; Length 1140;
Best Local Similarity 57.0%; Pred. No. 12;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Y 10 ccttgagatgagacaaagcagaagtcgtgaggaaggagccacacagtcaggaggag 69
Db 124 CTTGAGCTTGACATGTAATCTTCCAGCAATGCGCTGCGAACCACGAGCTCGAGGAGAG 65
Y 70 gtcccggtccagggcagctgctgag 95
Db 64 GAGCGGCTTCAGAGACCGGCGCGCG 39

RESULT 12
US-08-289-653-1/c
Sequence 1, Application US/08289653
Patent No. 5543322
GENERAL INFORMATION:
APPLICANT: Kazuaki KITANO et al.
TITLE OF INVENTION: DNA AND ITS USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,653
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: US/07/887,284
2 FILING DATE: May 22, 1992
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Warren M. Cheek, Jr.
5 REGISTRATION NUMBER: 33,367
6 REFERENCE/DOCKET NUMBER:
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: 202-371-8850
9 TELEFAX:
10
11
12 INFORMATION FOR SEQ ID NO: 1:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 2845 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: double
17 TOPOLOGY: linear
18
19 MOLECULE TYPE: Genomic DNA
20
21 HYPOTHETICAL:
22 ANTI-SENSE:
23 FRAGMENT TYPE:
24 ORIGINAL SOURCE:
25 ORGANISM: Fusarium sp.
26 STRAIN: S-19-5 (IFO 8084)
27 INDIVIDUAL ISOLATE:
28 DEVELOPMENTAL STAGE:
29 HAPLOTYPE:
30 TISSUE TYPE:
31 CELL TYPE:
32 CELL LINE:
33 ORGANELLE:
34 IMMEDIATE SOURCE:
35 LIBRARY:
36 CLONE:
37 POSITION IN GENOME:
38 CHROMOSOME/SEGMENT:
39 MAP POSITION:
40 UNITS:
41 FEATURE:
42 NAME/KEY:
43 LOCATION:
44 IDENTIFICATION METHOD:
45 OTHER INFORMATION:
46 PUBLICATION INFORMATION:
47 AUTHORS:
48 TITLE:
49 JOURNAL:
50 VOLUME:
51 ISSUE:
52 PAGES:
53 DATE:
54 DOCUMENT NUMBER:
55 FILING DATE:
56 PUBLICATION DATE:
57 RELEVANT RESIDUES IN SEQ ID NO:
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
9
```

```

1 STREET: 345 California Street
2 CITY: San Francisco
3 STATE: California
4 COUNTRY: USA
5 ZIP: 94104-2675
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: PatentIn Release #1.0, Version #1.30
12 CURRENT APPLICATION DATA:
13   APPLICATION NUMBER: US/08/164,292B
14 FILING DATE: 09-DEC-1993
15 CLASSIFICATION: 435
16 ATTORNEY/AGENT INFORMATION:
17 NAME: GRACEY, NANCY J
18 REGISTRATION NUMBER: 28,216
19 REFERENCE/DOCKET NUMBER: 29310-20021_00
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (415) 677-7000
22 TELEX: (415) 677-7522
23 FAX: 34-0154
24
25 INFORMATION FOR SEQ ID NO: 17:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 5100 base pairs
28 TYPE: nucleic acid
29 STRANDEDNESS: double
30 TOPOLOGY: linear
31 MOLECULE TYPE: DNA (genomic)
32 FEATURE:
33 NAME/KEY: CDS
34 LOCATION: 408..1331
35
36 US-08-164-292B-17
37
38 Query Match           26.5%; Score 26.8; DB 1; Length 5100;
39 Best Local Similarity 61.4%; Pred.No. 16;
40 Matches 43; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
41
42 Qy      9 gctcgtgaagacagaagtgcgtagtggaggaaaggccaccagtccaggaggaaa 68
43          | ||||| ||| || | |||| ||| ||| ||| ||| ||| ||| ||| |||
44 Db     1682 gtcttgagagcgacgccatcttcgctgcgcgaacagcacgcatctgcggcaa 1623
45          | ||||| ||| || | |||| ||| ||| ||| ||| ||| ||| ||| |||
46 Qy      69 gttccccgcgt 78
47          | ||||| |||
48 Db     1622 ctctcccctggt 1613
49
50 RESULT 15
51 US-08-164-292B-19/C
52 Sequence 19, Application US/08164292B
53 Patent No. 5820868
54 GENERAL INFORMATION:
55 APPLICANT: MITTAL, SURESH K.
56 APPLICANT: GRAHAM, FRANK L.
57 APPLICANT: PREVEC, LUDVIG
58 APPLICANT: BABIUK, LORENE A.
59 TITLE OF INVENTION: RECOMBINANT PROTEIN PRODUCTION IN BOVINE
60 TITLE OF INVENTION: ADENOVIRUS EXPRESSION VECTOR SYSTEM
61 NUMBER OF SEQUENCES: 34
62 CORRESPONDENCE ADDRESS:
63 ADDRESSEE: MORRISON & FOERSTER
64 STREET: 345 California Street
65 CITY: San Francisco
66 STATE: California
67 COUNTRY: USA
68 ZIP: 94104-2675
69
70 COMPUTER READABLE FORM:
71 MEDIUM TYPE: Floppy disk
72 COMPUTER: IBM PC compatible
73 OPERATING SYSTEM: PC-DOS/MS-DOS
74 SOFTWARE: PatentIn Release #1.0, Version #1.30
75 CURRENT APPLICATION DATA:
```

```

1  APPLICATION NUMBER: US/08/164,292B
2  FILING DATE: 09-DEC-1993
3  CLASSIFICATION: 435
4  ATTORNEY/AGENT INFORMATION:
5  NAME: GRACEY, NANCY J.
6  REGISTRATION NUMBER: 28,216
7  REFERENCE/DOCKET NUMBER: 29310-20021.00
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE: (415) 677-7000
10 TELEFAX: (415) 677-7522
11 TELEX: 34-0154
12
13 INFORMATION FOR SEQ ID NO: 19:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 5100 base pairs
16 TYPE: nucleic acid
17 STRANDEDNESS: double
18 TOPOLOGY: linear
19
20 MOLECULE TYPE: DNA (genomic)
21 FEATURE:
22 NAME/KEY: CDS
23 LOCATION: 529..954
24
25 US-08-164-292B-19
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
9
```

Search completed: October 3, 2002, 16:22:14
Job time: 16974 sec

Fri Oct 4 09:03:09 2002

us-09-826-581-5_copy_180_280.rni

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:49:24 ; Search time 7316.32 seconds

(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600

Perfect score: 101

Sequence: 1 cgtcccccgcagggcccatc.....atgcagagacacacctgcta 101

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| ----- | | | | | |

| | | | | | | |
|----|------|-------|--------|----|----------|----------------------|
| 1 | 101 | 100.0 | 989 | 6 | AX281579 | AX281579 Sequence |
| 2 | 101 | 100.0 | 1647 | 6 | AX281582 | AX281582 Sequence |
| 3 | 101 | 100.0 | 2109 | 6 | AX099776 | AX099776 Sequence |
| 4 | 101 | 100.0 | 2115 | 6 | AX099802 | AX099802 Sequence |
| 5 | 101 | 100.0 | 2115 | 9 | AF214519 | AF214519 Homo sapi |
| 6 | 101 | 100.0 | 152129 | 2 | AC027416 | AC027416 Homo sapi |
| 7 | 101 | 100.0 | 196554 | 2 | AC073128 | AC073128 Homo sapi |
| 8 | 97.8 | 96.8 | 2290 | 9 | HS249977 | HS249977 Homo sapi |
| 9 | 89.6 | 88.7 | 206854 | 9 | AC009974 | AC009974 Homo sapi |
| 10 | 72.4 | 71.7 | 1867 | 6 | AX099774 | AX099774 Sequence |
| 11 | 72.4 | 71.7 | 1873 | 4 | AF214520 | AF214520 Sus scrofa |
| 12 | 72.4 | 71.7 | 1873 | 6 | AX099800 | AX099800 Sequence |
| 13 | 72.4 | 71.7 | 2022 | 6 | AX099804 | AX099804 Sequence |
| 14 | 72.4 | 71.7 | 5888 | 4 | AF214521 | AF214521 Sus scrofa |
| 15 | 72.4 | 71.7 | 227724 | 2 | AF336381 | AF336381 Mus muscu |
| 16 | 33.4 | 33.1 | 3503 | 9 | HS181675 | HS181675 Homo sapien |
| 17 | 33.4 | 33.1 | 81811 | 2 | AC063981 | AC063981 Homo sapi |
| 18 | 33.4 | 33.1 | 82237 | 9 | AC003677 | AC003677 Human pl |
| 19 | 33.4 | 33.1 | 108123 | 9 | AC034206 | AC034206 Homo sapi |
| 20 | 33.4 | 33.1 | 159420 | 2 | AC009017 | AC009017 Homo sapi |
| 21 | 32 | 31.7 | 790 | 33 | AC076924 | AC076924 Giardia |
| 22 | 32 | 31.7 | 975 | 33 | AC078451 | AC078451 Giardia |
| 23 | 32 | 31.7 | 988 | 33 | AC028281 | AC028281 Giardia |
| 24 | 32 | 31.7 | 1927 | 3 | AF071896 | AF071896 Giardia |
| 25 | 31.2 | 30.9 | 17570 | 1 | AE007063 | AE007063 Mycobacte |
| 26 | 31.2 | 30.9 | 39430 | 1 | MTCV49 | MTCV49 Mycobacte |
| 27 | 30.4 | 30.1 | 736 | 33 | AC053985 | AC053985 Giardia |
| 28 | 30.4 | 30.1 | 804 | 33 | AC053984 | AC053984 Giardia |
| 29 | 30.4 | 30.1 | 967 | 33 | AC089076 | AC089076 Giardia |
| 30 | 30.4 | 30.1 | 1283 | 33 | AF050755 | AF050755 Giardia |
| 31 | 30.4 | 30.1 | 1306 | 3 | AF050754 | AF050754 Giardia |
| 32 | 30.4 | 30.1 | 3989 | 3 | AF071897 | AF071897 Giardia |
| 33 | 30.4 | 30.1 | 5791 | 1 | AE117694 | AE117694 Rhizobium |
| 34 | 30.2 | 29.9 | 11085 | 1 | AE004450 | AE004450 Pseudomon |
| 35 | 29.6 | 29.3 | 3603 | 1 | AF035395 | AF035395 Pseudomon |
| 36 | 29.6 | 29.3 | 3810 | 3 | AC091120 | AC091120 Leishman |
| 37 | 29.4 | 29.1 | 850 | 33 | AC042098 | AC042098 Giardia |
| 38 | 29.4 | 29.1 | 214609 | 2 | AC024651 | AC024651 Homo sapi |
| 39 | 29.2 | 28.9 | 56425 | 7 | AF165214 | AF165214 Bacteriop |
| 40 | 29 | 28.7 | 1575 | 3 | AY069691 | AY069691 Drosophil |
| 41 | 29 | 28.7 | 69061 | 2 | AC012986 | AC012986 Drosophil |
| 42 | 29 | 28.7 | 158628 | 2 | AC020653 | AC020653 Homo sapi |
| 43 | 29 | 28.7 | 168469 | 3 | AC007886 | AC007886 Drosophil |
| 44 | 29 | 28.7 | 173729 | 2 | AC012236 | AC012236 Homo sapi |
| 45 | 29 | 28.7 | 228448 | 3 | AE003772 | AE003772 Drosophil |

ALIGNMENTS

RESULT 1
AX281579 989 bp DNA linear PAT 02-NOV-2001
LOCUS Sequence 2 from Patent WO0177305.
DEFINITION AX281579
ACCESSION AX281579.1 GI:16608830
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 2 18-OCT-2001;
Arexis AB (SE)

FEATURES
source
1..989 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 229 a 306 c 286 g 168 t
ORIGIN

Query Match 100.0% Score 101: DB 6; Length 989;
Best Local Similarity 100.0%; Pred. No. 1.5e-17;
Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ctgtcccgagggcccatctcccaagctggctgagatgaagaaacccggc 60
|||||
DB 793 CTGTCCCGAGGCCCATTTTCCCAAGCTGGCTGCATGACGAACTCGGAAACCCGGC 852
|||||

OY 61 gccagatctacatgcgtctcatgcagagacacactgcta 101
|||||
DB 853 GCCAGATCTACATCGCTTCATGCAGACACACTGCTCTA 893
|||||

RESULT 2
AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Anderson, L., Luthman, H. and Marklund, S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 5 18-Oct-2001;
Atexis AB (SE)
FEATURES
source Location/Qualifiers
1..1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
20..1489
/note="unnamed protein product"
/codon_start=1
/protein_id="CA010589.1"
/db_xref="GI:16608834"
/translation="MEPGLLEHALRTPSMSSLOGSEHQMSEFLSENSSWSPAYTS
SSERIKGRRAKALRWTRKSVSEGEPRGPRPAESTGLEATPRTPLAQA
DPAGVGTPTGMDCLPSCDTASAAGSTDDVELATPEPTAMECELELBERPALC
LSPOAFEPKLGMDDELARKGAOITWRFMEHRCYDMATSKVITFDMLKKAFA
LVANGVRAAPKINDSKOSFVGMHTITDFTLVLRHRSPLVVOYELEQKIEFWREIF
LOGCFKPLVISISPNDSLEFVAVTLLKNRHLRLPVDPSGVNVLHITHRKLKFLHIF
GSLPRPSFLYRTIDYIGTFRDIAVLETAPIILALDIFDRLVSALPVNCEGOV
VGLYSREFDIHLAAOQTYNHLDMSEALRORTLCLEGLVSCOPHESLGEVIDRIARE
QVRLVLVDETOHLGCVSLSDIILQALVSPAGIDALGA"

BASE COUNT 346 a 502 c 462 g 337 t

ORIGIN

Query Match 100.0% Score 101: DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ctgtcccgagggcccatctcccaagctggctgagatgaagaaacccggc 60
|||||
DB 500 CTGTCCCGAGGCCCATTTTCCCAAGCTGGCTGCATGACGAAACCCGGC 559
|||||

OY 61 gccagatctacatgcgtctcatgcagagacacactgcta 101
|||||
DB 560 GCCAGATCTACATCGCTTCATGCAGACACACTGCTCTA 600
|||||

RESULT 3
AX099776 2109 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 3 from Patent WO0120003.
ACCESSION AX099776
VERSION AX099776.1 GI:13538810
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Anderson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Kogel-Galliard, C., Iannuccelli, N., Gellin, J., le Roy, P. and
Chardon, P.
TITLE Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof
JOURNAL Patent: WO 0120003-A 3 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Lelf (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES
source Location/Qualifiers
1..2109
/organism="Homo sapiens"
/db_xref="taxon:9606"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35799.1"
/db_xref="GI:13538811"
/translation="MRFMQEHRCYDMATSSKLVTFDTMLLEIKKAFALVANGVRAP
LMSKKOSFVGMHTITDFTLVLRHRSPLVVOYELEQKIEFWREIFLOGCFKPLVS
ISPNDSLFRAYVTTLLKNRHLRLPVDPSGVNVLHITHRKLKFLHIFGSLPRPSFL
YRTIDYIGTFRDIAVLETAPIILALDIFDRLVSALPVNCEGOVGLYSRDI
HLAAOQTYNHLDMSEALRORTLCLEGLVSCOPHESLGEVIDRIAREQVRLVLVDE
THLQGVSLSDIILQALVSPAGIDALGA"

BASE COUNT 458 a 621 c 560 g 470 t

ORIGIN

Query Match 100.0% Score 101: DB 6; Length 2109;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ctgtcccgagggcccatctcccaagctggctgagatgaagaaacccggc 60
|||||
DB 400 CTGTCCCGAGGCCCATTTTCCCAAGCTGGCTGCATGACGAAACCCGGC 459
|||||

OY 61 gccagatctacatgcgtctcatgcagagacacactgcta 101
|||||
DB 460 GCCAGATCTACATCGCTTCATGCAGACACACTGCTCTA 500
|||||

RESULT 4
AX099802 2115 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 29 from Patent WO0120003.
ACCESSION AX099802
VERSION AX099802.1 GI:13538836
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Anderson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Kogel-Galliard, C., Iannuccelli, N., Gellin, J., le Roy, P. and
Chardon, P.
TITLE Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof
JOURNAL Patent: WO 0120003-A 29 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Lelf (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES
source Location/Qualifiers
1..2115
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1395
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35801.1"
/db_xref="GI:13538837"

/translation="MSFLQENSSSWPSPAVTSSEERI RCKRRAKALRMTRQKSVVEG
EPPOGEGPRSPRTAESEGLATFPKTTPLAQQADPACVGPPTGMDCLPDSCTASAG
SSTDVLEATEPAPTEAMECELEGLLEERPLALCSPOAPFPKLGMDLRLKPGAQITM
RMOEHCTYDAMATSSKLVIFEDTMLLEIKAFALVANGVRAAPLMDSKOSFVGMILT
TDFILVLRHYRSPVVOIYEIOHKIETREIYLQGGFKPLVSPNDSJFEAVYTLI
KNRIHRLPVLDPVSGNVLIILTHRLKLEFLHIFCSLLRPSEFLRTIODGICFRL
AVVLETPILNALDI FVDRVSALPVNCEGVVGLSRDVIHLAQQYTNHLDMSV
GEALRORTLCLEGLVSCOPHSESLCEVIDRIAREQVHRLVLDVDETHLLGVVSLDILLO
ALVSPACIDALGAA"

BASE COUNT 460 a 622 c 562 g 471 t

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 2115;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgtcccgagagcccaattcccaagctggctggatgacgaactgagaaaccggc 60
|||||
DB 406 CTGTCCCGCAGGCCCATTTCCCAAGCTGGCTGGATGACGAACCTCGGAACCCGCC 465

QY 61 gccagatctacatgcgtctcatgcagagacacacctgcta 101
|||||
DB 466 GCCCAGATCTACATGCCCTTCATGCGAGCAGCACCTGCTA 506

RESULT 5
AF214519 2115 bp mRNA linear PRI 03-JUN-2000
LOCUS
DEFINITION Homo sapiens AMP-activated protein kinase gamma subunit (PRKG3)
ACCESSION AF214519
VERSION AF214519.1 GI:8215681
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2115)
Mellan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A., Thelander, M.,
Rogel-Gallard, C., Paul, S., Gellin, J., Lundstrom, K., Relensch, N.,
Lundstrom, K., Relensch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.
and Andersson, L.
A mutation in PRKG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS 2 (bases 1 to 2115)
Mellan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Gallard, C., Paul, S., Gellin, J., Lundstrom, K., Relensch, N.,
Kalm, E., Le Roy, P., Chardon, P. and Andersson, L.
Direct Submission
Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
Sweden

TITLE
JOURNAL

FEATURES
source Location/Qualifiers
1..2115
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p"
/tissue_type="skeletal muscle"
1..2115
/gene="PRKG3"
1..1395
/gene="PRKG3"
/note="AMPK3"
/note="AMPK3"
/codon_start=1
/product="AMP-activated protein kinase gamma subunit"
/protein_id="AAF73987.1"
/db_xref="GI:8215682"
/translation="MSFLQENSSSWPSPAVTSSSERIRCKRRAKALRMTRQKSVVEG

EPPOGEGPRSPRTAESEGLATFPKTTPLAQQADPACVGPPTGMDCLPDSCTASAG
SSTDVLEATEPAPTEAMECELEGLLEERPLALCSPOAPFPKLGMDLRLKPGAQITM
RMOEHCTYDAMATSSKLVIFEDTMLLEIKAFALVANGVRAAPLMDSKOSFVGMILT
TDFILVLRHYRSPVVOIYEIOHKIETREIYLQGGFKPLVSPNDSJFEAVYTLI
KNRIHRLPVLDPVSGNVLIILTHRLKLEFLHIFCSLLRPSEFLRTIODGICFRL
AVVLETPILNALDI FVDRVSALPVNCEGVVGLSRDVIHLAQQYTNHLDMSV
GEALRORTLCLEGLVSCOPHSESLCEVIDRIAREQVHRLVLDVDETHLLGVVSLDILLO
ALVSPACIDALGAA"

BASE COUNT 460 a 622 c 562 g 471 t

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 2115;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgtcccgagagcccaattcccaagctggctggatgacgaactgagaaaccggc 60
|||||
DB 406 CTGTCCCGCAGGCCCATTTCCCAAGCTGGCTGGATGACGAACCTCGGAACCCGCC 465

QY 61 gccagatctacatgcgtctcatgcagagacacacctgcta 101
|||||
DB 466 GCCCAGATCTACATGCCCTTCATGCGAGCAGCACCTGCTA 506

RESULT 6
AC027416 152129 bp DNA linear HTG 07-JUN-2000
LOCUS
DEFINITION Homo sapiens clone Rpl1-504G11, WORKING DRAFT SEQUENCE, 32
unordereed pieces.
AC027416
AC027416.2 GI:8317289
VERSION
HTG: HTGS_PHAISEL; HTGS_DRAFT.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 152129)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castelle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Lander, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talams, J.,
Testa, S., Theodore, J., Tirrell, A., Travers, M., Trifilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS 2 (bases 1 to 152129)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castelle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Lander, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talams, J.,
Testa, S., Theodore, J., Tirrell, A., Travers, M., Trifilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS 2 (bases 1 to 152129)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castelle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Lander, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talams, J.,
Testa, S., Theodore, J., Tirrell, A., Travers, M., Trifilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

COMMENT
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2000 this sequence version replaced gi:7342115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

```
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 17458
Center clone name: 504_C_11
-----
Summary Statistics
Sequencing vector: M13; M77815: 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135376 bases at least Q40
Consensus quality: 143264 bases at least Q30
Consensus quality: 146503 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 149029; sum-of-contigs
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 1006 1105: contig of 1005 bp in length
* 1106 2402: contig of 1297 bp in length
* 2403 2502: gap of 100 bp
* 2503 3823: contig of 1321 bp in length
* 3824 3923: gap of 100 bp
* 3924 5020: contig of 1097 bp in length
* 5021 5120: gap of 100 bp
* 5121 6161: contig of 1041 bp in length
* 6162 6261: gap of 100 bp
* 6262 7547: contig of 1286 bp in length
* 7548 7647: gap of 100 bp
* 7648 9983: contig of 2336 bp in length
* 9984 10083: gap of 100 bp
* 10084 12556: contig of 2473 bp in length
* 12557 12656: gap of 100 bp
* 12657 15043: contig of 2387 bp in length
* 15044 15143: gap of 100 bp
* 15144 17123: contig of 1980 bp in length
* 17124 17223: gap of 100 bp
* 17224 19466: contig of 2243 bp in length
* 19467 19566: gap of 100 bp
* 19567 21928: contig of 2362 bp in length
* 21929 22028: gap of 100 bp
* 22029 24319: contig of 2291 bp in length
* 24320 24419: gap of 100 bp
* 24420 27059: contig of 2640 bp in length
* 27060 27159: gap of 100 bp
* 27160 30170: contig of 3011 bp in length
* 30171 30270: gap of 100 bp
* 30271 33968: contig of 3698 bp in length
* 33969 34068: gap of 100 bp
* 34069 38179: contig of 4111 bp in length
* 38180 38279: gap of 100 bp
* 38280 42366: contig of 4087 bp in length
* 42367 42466: gap of 100 bp
* 42467 46365: contig of 3899 bp in length
* 46366 46465: gap of 100 bp
* 46466 51285: contig of 4820 bp in length
* 51286 51385: gap of 100 bp
* 51386 55871: contig of 4486 bp in length
* 55872 55971: gap of 100 bp
* 55972 60595: contig of 4624 bp in length
* 60596 60695: gap of 100 bp
* 60696 66595: contig of 5900 bp in length
* 66596 66695: gap of 100 bp
* 66696 73218: contig of 6523 bp in length
* 73219 73318: gap of 100 bp
-----
FEATURES
Source
* 73319 77115: contig of 3797 bp in length
* 77116 77215: gap of 100 bp
* 77216 85022: contig of 7807 bp in length
* 85023 85122: gap of 100 bp
* 85123 93314: contig of 8192 bp in length
* 93315 93414: gap of 100 bp
* 93415 101193: contig of 7779 bp in length
* 101194 101293: gap of 100 bp
* 101294 113090: contig of 11797 bp in length
* 113091 113190: gap of 100 bp
* 113191 123496: contig of 10306 bp in length
* 123497 123596: gap of 100 bp
* 123597 137837: contig of 14241 bp in length
* 137838 137937: gap of 100 bp
* 137938 152129: contig of 14192 bp in length.
Location/Qualifiers
1..152129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-504G11"
/clone_11b="RPCT-11 Human Male BAC"
1..1005
/note="assembly_fragment"
1106..2402
/note="assembly_fragment"
2503..3823
/note="assembly_fragment"
3924..5020
/note="assembly_fragment"
5121..6161
/note="assembly_fragment"
6262..7547
/note="assembly_fragment"
7648..9983
/note="assembly_fragment"
10084..12556
/note="assembly_fragment"
12657..15043
/note="assembly_fragment"
15144..17123
/note="assembly_fragment"
17224..19466
/note="assembly_fragment"
19567..21928
vector_side:left
22029..24319
/note="assembly_fragment"
24420..27059
/note="assembly_fragment"
27160..30170
/note="assembly_fragment"
30271..33968
/note="assembly_fragment"
34069..38179
/note="assembly_fragment"
38280..42366
/note="assembly_fragment"
42467..46365
/note="assembly_fragment"
46466..51285
/note="assembly_fragment"
51386..55871
/note="assembly_fragment"
55972..60595
/note="assembly_fragment"
60696..66595
/note="assembly_fragment"
66696..73218
/note="assembly_fragment"
73319..77115
/note="assembly_fragment"
```

```

misc_feature      77216..85022
                  /note="assembly_fragment"
misc_feature      85123..93314
                  /note="assembly_fragment"
misc_feature      93415..101193
                  /note="assembly_fragment"
misc_feature      101294..113090
                  /note="assembly_fragment"

Query Match      100.0%; Score 101; DB 2; Length 152129;
Host Local Similarity 100.0%; Pred. No. 1e-17;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      ctgtcccgagggcccatcttcccaagctggcgtggaatgacgaactgcgaaaccgcgc 60
db      38135  ctgtcccgagggcccatcttcccaagctggcgtggaatgacgaactgcgaaaccgcgc 38076

Qy      61      gccacatcatatcgctcatcagagagaccacticta 101
db      38075  gccacatctacatgcgcttccatgcacagcacaccttcta 38035

RESULT 7
AC073128/c      196554 bp      DNA      linear      HTG 21-FEB-2001
LOCUS      Homo sapiens chromosome 2 clone RP11-64705, WORKING DRAFT SEQUENCE,
DEFINITION      17 unordered pieces.
ACCESSION      AC073128
VERSION      AC073128.3 GI:13027579
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 196554)
AUTHORS      Waterston,R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 196554)
TITLE      Waterston,R.H.
JOURNAL      Submitted (08-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT      On Feb 21, 2001 this sequence version replaced gi:8469048.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information -----
Center project name: H.NH0647005
Summary Statistics -----
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator ET; 98% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187795 bases at least Q40
Consensus quality: 190513 bases at least Q30
Consensus quality: 192099 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 194954; sum-of-contigs
Quality coverage: 5.58 in Q20 bases; agarose-fp
Quality coverage: 5.67 in Q20 bases; sum-of-contigs
----- NOTE -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved.
1      1157: contig of 1157 bp in length
1158      1257: gap of unknown length
1258      3600: contig of 2143 bp in length
3601      3700: gap of unknown length
3701      5103: contig of 1403 bp in length
5104      5203: gap of unknown length
5204      8524: contig of 3121 bp in length
8525      8624: gap of unknown length
8625      11856: contig of 3232 bp in length
11857      11956: gap of unknown length
11957      15783: contig of 3827 bp in length
15784      15883: gap of unknown length
15884      21906: contig of 6023 bp in length
21907      22006: gap of unknown length
22007      28887: contig of 6881 bp in length
28888      28987: gap of unknown length
28988      35255: contig of 6268 bp in length
35256      35355: gap of unknown length
35356      44642: contig of 9287 bp in length
44643      44742: gap of unknown length
44743      58275: contig of 13533 bp in length
58276      58375: gap of unknown length
58376      73816: contig of 15441 bp in length
73817      73916: gap of unknown length
73917      92140: contig of 18224 bp in length
92141      92240: gap of unknown length
92241      113337: contig of 21097 bp in length
113338      113437: gap of unknown length
113438      130325: contig of 16888 bp in length
130326      130425: gap of unknown length
130426      149287: contig of 18862 bp in length
149288      149388: gap of unknown length
149389      196554: contig of 47167 bp in length.

FEATURES
source      Location/Qualifiers
1..196554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-64705"
1..1157
/note="assembly_name:Contig17"
1158..3600
/note="assembly_name:Contig18"
3701..5103
/note="assembly_name:Contig19
vector_end:77
vector_side:right"
5204..8524
/note="assembly_name:Contig20"
8625..11856
/note="assembly_name:Contig21"
11957..15783
/note="assembly_name:Contig22"
15884..21906
/note="assembly_name:Contig23"
22007..28887
/note="assembly_name:Contig24"
28988..35255
/note="assembly_name:Contig25"
35356..44642
/note="assembly_name:Contig26"
44743..58275
/note="assembly_name:Contig27"
58376..73816
/note="assembly_name:Contig28"
73917..92140
/note="assembly_name:Contig29"
92241..113337
/note="assembly_name:Contig30"
113438..130325
/note="assembly_name:Contig31"
130426..149287

```

```

misc_feature      /note="assembly_name:Contig32"
                  149388..196554
                  /note="assembly_name:Contig33
                  clone_end:SP6
                  vector_side:right"
BASE COUNT      52296 a 46993 c 45889 g 49770 t 1606 others
ORIGIN
Query Match      100.0%: Score 101; DB 2; Length 196554;
Best Local Similarity 100.0%: Pred. No. 1e-17;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 cgtgccgcagagcccatcctcccaagctggctggatgacgaactcgcgaacccgcg 60
Db 62634 CTGTCCCGCAGGCCCATTTCCCAAGCTGGCTGGATGACGAACCTCGGAACCCGCC 62575
Oy 61 gccagatctacatgccttcatgcaagagcacactgcta 101
Db 62574 GCCCAGATCTACATCGCTTCATCGAGACACACTCTCTA 62534

RESULT 8
LOCUS      HSA249977      2290 bp      mRNA      linear      PRI 07-APR-2000
DEFINITION      Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
ACCESSION      AJ249977
VERSION      AJ249977.1 GI:6688200
KEYWORDS      AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2290)
Cheung, P.C., Salt, I.P., Davies, S.P., Hardie, D.G. and Carling, D.
Characterization of AMP-activated protein kinase gamma-subunit
isoforms and their role in AMP binding
Biochem. J. 346 Pt 3, 659-669 (2000)
20164049
2 (bases 1 to 2290)
Carling, D.
Direct Submission
Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
London, W12 0NN, UNITED KINGDOM
Location/Qualifiers
1..2290
/organism="Homo sapiens"
/db_xref="taxon:9606"
22..1500
/gene="AMPK gamma 3"
22..1500
/gene="AMPK gamma 3"
22..1500
/gene="AMPK gamma 3"
/function="AMP-activated protein kinase regulatory
subunit"
/codon_start=1
/evidence-experimental
/product="AMP-activated protein kinase gamma 3 subunit"
/protein_id="CA865117.1"
/db_xref="GI:6688201"
/translation="MEPGLHALRRTPSMSSLSGSEHOEMSPLEOENSSSPSPAYTS
SSERIRGRRAKALRMTRKSVSEGGPGGEGRRRAESNGIETKTPKTPPLQA
DPAGVPTPTMDCLPSDCTASAGSSTIDVELATPEPRAEAGCELEGLERPALC
LSPOAPFKLGMDELRRKPGAOIYMRIFERHTCYDMATSSKLVITDMLETKKAFEA
LVANGVRAAPLMDSKKSPFGMLITPTFIVLHRYRSPVLVLEYIEQHKITKREIF
LGGFKPLVSTSPNDLSFEAVYTLIKRRIHRLPVLDPVSGNVLAILTHRRLKFLHIF
GSLPRPSPLKRTIQQDGLGTGFNDLAVLETADILTDALDIFVDRVSALPVNVECGOV
VGLSRFDVTHLAQDTYNHLDMSVGEALRKRLCLEGVISCPHSLGELVIDRLARE
OVHRLVIVDEFQHLGVSLSDILOALVISPACIDPSGPEKI"
BASE COUNT      501 a 674 c 617 g 498 t
ORIGIN

```

```

Query Match      96.8%: Score 97.8; DB 9; Length 2290;
Best Local Similarity 98.0%: Pred. No. 1e-16;
Matches 99; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 cgtgccgcagagcccatcctcccaagctggctggatgacgaactcgcgaacccgcg 60
Db 502 CTGTCCCGCAGGCCCATTTCCCAAGCTGGCTGGATGACGAACCTCGGAACCCGCC 561
Oy 61 gccagatctacatgccttcatgcaagagcacactgcta 101
Db 562 GCCCAGATCTACATCGCTTCATCGAGACACACTCTCTA 602

RESULT 9
LOCUS      AC009974      206854 bp      DNA      linear      PRI 09-JAN-2002
DEFINITION      Homo sapiens BAC clone RP11-459119 from 2, complete sequence.
ACCESSION      AC009974
VERSION      AC009974.9 GI:16799058
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 206854)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1996)
99063792
2 (bases 1 to 206854)
Harris, A. and Cotton, M.
The sequence of Homo sapiens BAC clone RP11-459119
Unpublished (2001)
3 (bases 1 to 206854)
Waterston, R.H.
Direct Submission
Submitted (08-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 206854)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 206854)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 8 2001 this sequence version replaced g1:13431203.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
-----
Summary Statistics
Center project name: H_NH0459119
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenoe, M., Calanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-1077K22; the clone sequenced to the right is RP11-64705. Actual start of this clone is at base position 1 of RP11-459119; actual end is at base position 206854 of RP11-459119.

Data from AC079810 and AC073128 was used to finish this clone. AC009974. Polymorphisms have been identified between AC073128 and AC009974. A single plasmid region exists between 38812-38903. An unresolved tandem in the HERY SVA exists between 184390-185163. PCR suggests that approximately 1700 bps are missing.

Location/Qualifiers

1..206854

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone_11b="RPCI-11"

misc_feature

/note="match to EST BE314060 (NID:g1347719)"

misc_feature

/note="match to EST BI059713 (NID:g14467240)"

misc_feature

/note="match to EST BF183086 (NID:g11061273)"

misc_feature

/note="match to EST AL567345 (NID:g12920610)"

misc_feature

/note="match to EST AW880850 (NID:g8042860)"

misc_feature

/note="match to EST BF304755 (NID:g11251653)"

misc_feature

/note="similar to Homo sapiens EST B1114348 (NID:g14565249)"

misc_feature

/note="match to EST BG477625 (NID:g13409904)"

misc_feature

/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature

/note="match to EST BE908408 (NID:g10402954)"

misc_feature

/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

misc_feature

/note="match to EST C05773 (NID:g1502549)"

misc_feature

/note="match to EST BG470047 (NID:g13402322)"

misc_feature

/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature

/note="match to EST BE908408 (NID:g10402954)"

misc_feature

281..344
/note="match to EST BF304755 (NID:g11251653)"

misc_feature

281..344
/note="match to EST BF183086 (NID:g11061273)"

misc_feature

281..344
/note="similar to Homo sapiens EST B1114348 (NID:g14565249)"

misc_feature

281..344
/note="match to EST BG477625 (NID:g13409904)"

misc_feature

281..344
/note="match to EST BI059713 (NID:g14467240)"

misc_feature

281..344
/note="match to EST AL567345 (NID:g12920610)"

misc_feature

281..344
/note="match to EST BE314060 (NID:g9134719)"

misc_feature

281..344
/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

misc_feature

281..344
/note="match to EST AW880850 (NID:g8042860)"

misc_feature

294..344
/note="match to EST BG470047 (NID:g13402322)"

misc_feature

complement(557..558)
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"

misc_feature

588..929
/note="match to EST T64073 (NID:g667938) yc05d12.r1"

misc_feature

594..763
/note="match to EST BF304755 (NID:g11251653)"

misc_feature

594..763
/note="match to EST BG470047 (NID:g13402322)"

misc_feature

594..763
/note="match to EST AW880850 (NID:g8042860)"

misc_feature

594..731
/note="match to EST BE314060 (NID:g9134719)"

misc_feature

594..764
/note="match to EST BI059713 (NID:g14467240)"

misc_feature

594..763
/note="match to EST BG477625 (NID:g13409904)"

misc_feature

594..742
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"

misc_feature

594..763
/note="match to EST BE908408 (NID:g10402954)"

misc_feature

594..763
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"

misc_feature

594..763
/note="match to EST A1670836 (NID:g4850567) wa04g10.x1"

misc_feature

594..763
/note="match to EST BF183086 (NID:g11061273)"

misc_feature

594..763
/note="similar to Homo sapiens EST B1114348 (NID:g14565249)"

misc_feature

599..763
/note="match to EST AL567345 (NID:g12920610)"

misc_feature

617..1084
/note="match to EST AA481361 (NID:g2210913) zv44e01.r1"

misc_feature

622..763
/note="match to EST A1860958 (NID:g5514574) w156f05.x1"

misc_feature

622..763
/note="similar to Mus musculus EST A1196847 (NID:g3749453) u167e01.x1"

misc_feature

684..763
/note="match to EST BG992568 (NID:g14396638)"

misc_feature

962..1084
/note="match to EST A1656812 (NID:g4740791) ct54b06.x1"

misc_feature

967..1084
/note="match to EST BE908408 (NID:g10402954)"

misc_feature

967..1085
/note="match to EST BF304755 (NID:g11251653)"

misc_feature

967..1091
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"

misc_feature

Query Match 88.7%: Score 89.6; DB 9: Length 206854;
Best Local Similarity 95.8%: Pred. No. 1.2e-14;

Matches 92: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

Oy 6 cccgcagagcccatcctcccaagctggcgtgagatgacgaactcggaaccggcgccca 65

Db 168209 CCCCAGAGCCCATTTTCCCAAGCTGGGCTGGATGACGAACTGCGAAACCGCGGCCCA 168150

Oy 66 gacatcatgctgcttcatgacgaagacacacctgcta 101

Db 168149 GATCTACATGCGCTTCATGACGACACACCTGCTTA 168114

RESULT 10

AX099774

LOCUS AX099774 1867 bp DNA Linear PAT 02-APR-2001
DEFINITION Sequence 1 from Patent WO0120003.
ACCESSION AX099774
VERSION AX099774.1 GI:13538808
KEYWORDS
SOURCE
ORGANISM

pig. *Sus scrofa*

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
Andersson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Rogel-Galliard, C., Iannucci, N., Gellin, J., le Roy, P., and
Chardon, P.

REFERENCE

AUTHORS

TITLE

Variant of the gamma chain of amyl, dna sequences encoding the
same, and uses thereof
Patent: WO 0120003-A 1 22-MAR-2001:
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Lell (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)

Location/Qualifiers
1..1867
/organism="Sus scrofa"
/db_xref="taxon:9823"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35798.1"
/db_xref="GI:13538809"
/translation="MHFMQEHCTCYDAMATSSKLVIFDTMLIKKAFALVANGVRAAP
LMDSKKOSFVGMITIDFLVLRHYRSPLOYIEERKLETREIYLGCCFRPLVS
ISPNDSLFEAVVALIKNRIHRLPVLDPVSGAVLILIKRLIKLHITGLIPRPSFL
YRIODLGIGTFRLAVVLETPALIDIPDRVSAALPVNENETGOVGLYSRFDVI
HLAAGOTVNHLDNMGCEALROTTLCLEGVLSQPHETLGEVIDRIVREOVHRLVLVDE
TQHLGVVSLSDIILALVLSFAGIDALGA"

BASE COUNT 380 a 583 c 529 g 375 t

CDS

Query Match 71.7%: Score 72.4: DB 6: Length 1867;
Best Local Similarity 83.7%: Pred. No. 8, 2e-10;
Matches 82: Conservative 0: Mismatches 16: Indels 0: Gaps 0;

Oy 4 tccgcagagcccatcctcccaagctggcgtgagatgacgaactcggaaccggcgcc 63

Db 403 TCCCGAGAGCTGCTGTACCCAGCGCTGGATGATGACCTGCAGAACCGGGGCC 462

Oy 64 cagatcatgctgcttcatgacgaagacacacctgcta 101

Db 463 CAGCTCTACATGCTTCATGACGACACACCTGCTTA 500

RESULT 11

AF214520 1873 bp mRNA linear MAM 03-JUN-2000

LOCUS AF214520 1873 bp mRNA linear MAM 03-JUN-2000

DEFINITION Sus scrofa AMP-activated protein kinase gamma subunit (PRAG3)

ACCESSION AF214520

VERSION AF214520.1 GI:8215683

KEYWORDS

SOURCE pig. *Sus scrofa*

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
20280150
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A., Thelander, M.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

JOURNAL
MEDLINE
PUBMED
10818001
2 (bases 1 to 1873)
Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A.,
Rogel-Galliard, C., Paul, S., Iannucci, N., Raak, L., Ronne, H.,
Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P.L., Chardon, P.,
and Andersson, L.
A mutation in PRAG3 associated with excess glycogen content in pig
skeletal muscle
Science 288 (5469), 1248-1251 (2000)

REFERENCE
AUTHORS
TITLE

Search completed: October 3, 2002, 14:49:54
Job time: 12114 sec

```
Query Match      71.7%: Score 72.4; DB 4; Length 5888;
Best Local Similarity 83.7%: Pred. No. 7.6e-10;
Matches 82: Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 4 tccccgagggcccaattcccaagctgggtggtgagatgacgaactgcggaacccggcgcc 63
      ||||| ||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 TCCCCAGAGGTGCTGTTACCCAGGCTGGCTGGATGATGAGCTGCAGAGCCGGGGCCC 828
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 cagatctacatgcatcaltcagatgcagagacacactgcta 101
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 CAGGCTACATGCACCTTCATGCAGAGACACACCTGCTA 866

RESULT 15
AF336381/c 227724 bp DNA linear HTG 02-APR-2001
LOCUS Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN
DEFINITION
ACCESSION AF336381
VERSION AF336381.1 GI:13507298
KEYWORDS HTG: HTGS_PBASE1.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 227724)
JOURNAL Rump, A., Hayes, C., Brown, S.D.M. and Rosenthal, A.
REFERENCE Mouse chromosome 1 genomic sequence
AUTHORS Unpublished
JOURNAL 2 (bases 1 to 227724)
Rump, A.
Direct Submission
Submitted (17-JAN-2001) Genome Analysis Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT * NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 17869: contig of 17869 bp in length
* 17870 17969: gap of unknown length
* 17970 32746: contig of 14777 bp in length
* 32747 32846: gap of unknown length
* 32847 227724: contig of 194878 bp in length.
FEATURES
Source
Location/Qualifiers
1..227724
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/clone="PAC510; PAC457"
BASE COUNT 57663 a 55423 c 56238 g 58199 t 201 others
ORIGIN
```

```
Query Match      71.7%: Score 72.4; DB 2; Length 227724;
Best Local Similarity 83.7%: Pred. No. 5.8e-10;
Matches 82: Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 4 tccccgagggcccaattcccaagctgggtggtgagatgacgaactgcggaacccggcgcc 63
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8211 TCCCCAGAGGCCCCACTCTGCTGCGCTGAGTGGATGATGAGAACTTCAGAAACCCGGAGGCC 8152
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 cagatctacatgcatcaltcagatgcagagacacactgcta 101
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8151 CAGGCTACATGCACCTTCATGCAGAGACACACCTGTTA 8114
```

FT XX /note= "5' portion of intron 4"
PN MO200177305-A2.
XX 18-OCT-2001.
PD 06-APR-2001; 2001WO-SE00765.
XX 07-APR-2000; 2000US-195665P.
PR (AREX-) AREXIS AB.
XX
PA Andersson L, Luthman H, Marklund S;
PI WPI: 2001-657170/75.
DR
XX
XX
PT New variants of human AMP-activated protein kinase gamma3 subunit
PT associated with a metabolic disease e.g. diabetes or obesity and method
PT for determining a risk estimate of diseases in subject by detecting the
PT variant -
XX
PS Example 1; Fig 2; 25pp; English.
XX
CC The sequences given in AAH43681-84 represents genomic fragments
CC encoding the human AMP-activated protein kinase gamma 3 subunit
CC (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC is useful in determining a risk estimate of a metabolic disease,
CC such as diabetes or obesity, in a subject. The variation may occur
CC in exons 3, 4 or 10. In exon 3 variation may be a substitution of
CC a G for a C at nucleotide 320, resulting in the amino acid
CC substitution P71A; in exon 4 variation may be a substitution of a
CC T for a C at nucleotide 550; and in exon 10 variation may be a
CC substitution of a T for a C at nucleotide 1037, resulting in the
CC amino acid substitution R340W. There may also be nucleotide variation
CC in Intron 6.
XX
SQ Sequence 989 BP; 229 A; 306 C; 286 G; 168 T; 0 other;
XX
Query Match 100.0%; Score 101; DB 22; Length 989;
Best Local Similarity 100.0%; Pred. No. 9.8e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ctgtcccgagagcccaattcccaagctggctggatgacgaactgcggaaccggc 60
DB 793 ctgtcccgagagcccaattcccaagctggctggatgacgaactgcggaaccggc 852
OY 61 gccagatctacatgcgtctcatgcagagcacacctgcta 101
DB 853 gccagatctacatgcgtctcatgcagagcacacctgcta 893
RESULT 2
AAH43685
ID AAH43685 standard; cDNA; 1647 BP.
XX
AC AAH43685;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 cDNA.
XX
KM Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KM metabolic disease: diabetes; obesity; substitution; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 20..1489
FT FT /*tag= a
FT FT /product= "PRKAG3"
FT FT 230
FT FT /*tag= b
FT FT variation

FT FT /label= "C230G"
FT FT /note= "Causes P71A"
FT FT 559
FT variation
FT FT /*tag= C
FT FT /label= "T559C"
FT FT /note= "Silent variation"
FT FT 1037
FT variation
FT FT /*tag= d
FT FT /label= "C1037T"
FT FT /note= "Causes R340W"
XX
PN MO200177305-A2.
XX
XX 18-OCT-2001.
PD 06-APR-2001; 2001WO-SE00765.
XX 07-APR-2000; 2000US-195665P.
PR (AREX-) AREXIS AB.
XX
PA Andersson L, Luthman H, Marklund S;
PI WPI: 2001-657170/75.
DR P-PSDB: QDB47679.
DR
XX
XX
PT New variants of human AMP-activated protein kinase gamma3 subunit
PT associated with a metabolic disease e.g. diabetes or obesity and method
PT for determining a risk estimate of diseases in subject by detecting the
PT variant -
XX
PS Disclosure: Fig 5; 25pp; English.
XX
XX This sequence represents the full length cDNA encoding the human
CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC a risk estimate of a metabolic disease, such as diabetes or obesity,
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon
CC 3 variation may be a substitution of a G for a C at nucleotide 320,
CC resulting in the amino acid substitution P71A; in exon 4 variation may
CC be a substitution of a T for a C at nucleotide 550; and in exon 10
CC variation may be a substitution of a T for a C at nucleotide 1037,
CC resulting in the amino acid substitution R340W. There may also be
CC nucleotide variation in Intron 6. The numbering of these
CC variations is based on the full length cDNA as given, rather than on
CC position 1 of the open reading frame.
XX
SQ Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;
XX
Query Match 100.0%; Score 101; DB 22; Length 1647;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ctgtcccgagagcccaattcccaagctggctggatgacgaactgcggaaccggc 60
DB 500 ctgtcccgagagcccaattcccaagctggctggatgacgaactgcggaaccggc 559
OY 61 gccagatctacatgcgtctcatgcagagcacacctgcta 101
DB 560 gccagatctacatgcgtctcatgcagagcacacctgcta 600
RESULT 3
AAD03296
ID AAD03296 standard; DNA; 2109 BP.
XX
AC AAD03296;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX

KW human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
 KW cystathione beta synthase; CBS; cardiact; gene therapy; ss.
 OS Homo sapiens.
 FI Key Location/Qualifiers
 FT 5'UTR 1..471
 FT CDS /*tag= a
 FT 472..1389
 FT /*tag= b
 FT /product= "Human Prkag3 protein"
 FT 1390..2109
 FT /*tag= c
 PN WO200120003-A2.
 PD 22-MAR-2001.
 PF 11-SEP-2000: 2000MO-EP09896.
 PR 10-SEP-1999: 99EP-0402236.
 PR 18-MAY-2000: 2000EP-0401388.
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA (ANDE/) ANDERSSON L.
 PA (LOOF/) LOOFT C.
 PA (KALM/) KALM E.
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Galliard C;
 PI Iannucciell N, Gellin J, Le Roy P, Chardon P;
 DR MPI: 2001-244810/25.
 DR P-PSDB: AAE00221.
 XX New variants of the gamma subunit of vertebrate adenosine
 PT monophosphate-activated kinase for diagnosis or treatment of disorders
 PT associated with energy metabolism such as diabetes, obesity, and
 PT myopathy -
 PS Claim 12: Fig 2: 71pp: English.
 XX The present sequence is a cDNA encoding human adenosine monophosphate
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
 CC PRKAG3. Mutation in Prkag3 results in an altered regulation of
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
 CC useful as therapeutic for treating carbohydrate metabolism disorders such
 CC as diabetes, obesity, and disorders associated with muscle metabolism
 CC such as myopathy and cardiovascular diseases, to modulate AMPK
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
 CC and its functionally altered mutants are useful for the diagnostic
 CC evaluation, genetic testing and prognosis of a metabolic disorder,
 CC preferably a carbohydrate metabolism disorder. Primers that can detect
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting
 CC from the expression of a functionally altered allele of PRKAG3.
 CC Transgenic animal and host cell transformed with PRKAG3 or a
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
 CC screening compounds able to modulate AMPK activity. Nucleic acid
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain
 CC of PRKAG3 and is useful in gene therapy.
 XX Sequence 2109 BP: 458 A; 621 C; 560 G; 470 T; 0 other:
 SQ
 Query Match 100.0%; Score 101; DB 22; Length 2109;
 Best Local Similarity 100.0%; Pred. NO. 1..1e-22;
 Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0y 1 cgtgcgccgagccgcatctcccaagctggtggtgagtcgcaactcgcgaaccgcgc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 400 ctgcccgcgagggcccatctcccaagctggtggtgagtcgcaactcgcgaaccgcgc 459
 0y 61 gcccaagctacatgcgtcttcacgagagacacactgcta 101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 460 gcccaagctacatgcgtcttcacgagagacacactgcta 500
 RESULT 4
 AAD03320
 ID AAD03320 standard; cDNA: 2115 BP.
 XX
 AC AAD03320;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
 XX
 KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
 KW cystathione beta synthase; CBS; cardiact; gene therapy; ss.
 XX
 OS Homo sapiens.
 FI Key Location/Qualifiers
 FT CDS 1..1395
 FT /*tag= a
 FT /product= "Human complete Prkag3 protein"
 XX
 PN WO200120003-A2.
 PD 22-MAR-2001.
 PF 11-SEP-2000: 2000MO-EP09896.
 PR 10-SEP-1999: 99EP-0402236.
 PR 18-MAY-2000: 2000EP-0401388.
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA (ANDE/) ANDERSSON L.
 PA (LOOF/) LOOFT C.
 PA (KALM/) KALM E.
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Galliard C;
 PI Iannucciell N, Gellin J, Le Roy P, Chardon P;
 DR MPI: 2001-244810/25.
 DR P-PSDB: AAE00223.
 XX New variants of the gamma subunit of vertebrate adenosine
 PT monophosphate-activated kinase for diagnosis or treatment of disorders
 PT associated with energy metabolism such as diabetes, obesity, and
 PT myopathy -
 PS Claim 12: Page 65-68: 71pp: English.
 XX The present sequence is a cDNA encoding human adenosine monophosphate
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
 CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
 CC useful as therapeutic for treating carbohydrate metabolism disorders such
 CC as diabetes, obesity, and disorders associated with muscle metabolism
 CC such as myopathy and cardiovascular diseases, to modulate AMPK
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
 CC and its functionally altered mutants are useful for the diagnostic
 CC evaluation, genetic testing and prognosis of a metabolic disorder,
 CC preferably a carbohydrate metabolism disorder. Primers that can detect
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting
 CC from the expression of a functionally altered allele of PRKAG3.
 CC Transgenic animal and host cell transformed with PRKAG3 or a
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
 CC screening compounds able to modulate AMPK activity. Nucleic acid

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathionine beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.

SO Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 101; DB 22; Length 2115;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ctgtcccgagagcccatctcccaagctggctggatgacgaactgcggaaccggc 60
Db 406 ctgtcccgagagcccatctcccaagctggctggatgacgaactgcggaaccggc 60
Oy 61 gccagatctacatgcgcttcacatgcagagacacactgcta 101
Db 466 gccagatctacatgcgcttcacatgcagagacacactgcta 506

RESULT 5

AAD03295
ID AAD03295 standard; cDNA; 1867 BP.

AC AAD03295;

XX 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.

XX Pig: gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KM PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathionine beta synthase; CBS; cardiact; gene therapy; RN locus;
XX chromosome 15; ss.

OS Sus scrofa.

XX Key Location/Qualifiers

FT 5'UTR 1..471

FT CDS /*tag= a

FT /*tag= b

FT /*product= "Sus scrofa PRKAG3 protein"

FT 3'UTR 1390..1867

FT /*tag= c

XX WO200120003-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP09896.

XX 10-SEP-1999; 99EP-0402236.

XX 18-MAY-2000; 2000EP-0401388.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX (ANDE/) ANDERSSON L.

XX (LOOF/) LOOFT C.

XX (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Mijan D, Robic A, Rogel-Gaillard C;
XX Iannucci N, Gellin J, Le Roy P, Chardon P;

XX WPI: 2001-244810/25.

XX P-PSDB: AAE00220.

XX New variants of the gamma subunit of vertebrate adenosine

XX monophosphate-activated kinase for diagnosis or treatment of disorders

XX myopathy associated with energy metabolism such as diabetes, obesity, and

XX Claim 12; Fig 2; 71pp; English.

XX The present sequence is a cDNA encoding pig adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder.
CC Preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathionine beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.

SO Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;

Query Match

Best Local Similarity 71.7%; Score 72.4; DB 22; Length 1867;
Matches 82; Conservative 83.7%; Pred. No. 1.3e-13; Mismatches 16; Indels 0; Gaps 0;

Oy 4 tccccgagagcccatctcccaagctggctggatgacgaactgcggaaccggc 63
Db 403 tccccgagagctgtcgtaccagcagctggctggatgacgaactgcggaaccggc 63
Oy 64 cagatctacatgcgcttcacatgcagagacacactgcta 101
Db 463 cagatctacatgcgcttcacatgcagagacacactgcta 500

RESULT 6

AAD03319
ID AAD03319 standard; cDNA; 1873 BP.

AC AAD03319;

XX 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.

XX Pig: gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KM PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathionine beta synthase; CBS; cardiact; gene therapy; RN locus;
XX chromosome 15; ss.

OS Sus scrofa.

XX Key Location/Qualifiers

FT CDS /*tag= a

FT /*tag= b

FT /*product= "Sus scrofa complete Prkag3 protein"

XX WO200120003-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP09896.

XX 10-SEP-1999; 99EP-0402236.

XX 18-MAY-2000; 2000EP-0401388.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX (ANDE/) ANDERSSON L.

| | |
|-----------------------|--|
| PA | (LDOOF//LDOOFT C. |
| PB | (KALM//) KALM E. |
| XX | |
| PI | Andersson L, Looft C, Kalm F, Milan D, Robic A, Rogel-Galliard G; |
| PJ | Iannuccelli N, Gellin J, Le Roy P, Chardon P; |
| DR | WPI: 2001-244810/725. |
| XX | |
| P-PSDB: | AAE00222. |
| PT | New variants of the gamma subunit of vertebrate adenosine |
| PR | monophosphate-activated kinase for diagnosis or treatment of disorders |
| PT | associated with energy metabolism such as diabetes, obesity, and |
| XX | myopathy - |
| PS | |
| PM | Claum 12; Page 62-64; 71pp: English. |
| XX | |
| CC | The present sequence is a cDNA encoding pig adenosine monophosphate |
| CC | (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform, |
| CC | complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome |
| CC | 15. Mutation in prkag3 results in an altered regulation of carbohydrate |
| CC | metabolism, particularly in skeletal muscle. PRKAG3 is useful as |
| CC | therapeutic for treating carbohydrate metabolism disorders such as |
| CC | diabetes, obesity, and disorders associated with muscle metabolism |
| CC | such as myopathy and cardiovascular diseases, to modulate AMPK |
| CC | activity, and for restoring a normal AMPK function. PRKAG3 sequence |
| CC | and its functionally altered mutants are useful for the diagnostic |
| CC | evaluation, genetic testing and prognosis of a metabolic disorder, |
| CC | preferably a carbohydrate metabolism disorder. Primers that can detect |
| CC | a genetic polymorphic marker linked to a sequence encoding PRKAG3, are |
| CC | useful for detecting a dysfunction of carbohydrate metabolism resulting |
| CC | from the expression of a functionally altered allele of PRKAG3. |
| CC | Transgenic animal and host cell transformed with PRKAG3 or a |
| CC | heterocimeric AMPK consisting of PRKAG3 or its mutant, are useful for |
| CC | screening compounds able to modulate AMPK activity. Nucleic acid |
| CC | encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or |
| CC | in a sequence encoding the first cystathione beta synthase (CBS) domain |
| CC | of PRKAG3 and is useful in gene therapy. |
| XX | |
| XQ | Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other; |
| XX | |
| Query Match | 71.7%; Score 72.4; DB 22; Length 1873; |
| Best Local Similarity | 83.7%; Pred. No. 1,3e-13; |
| Matches | 82; Conservative 0; Mismatches 16; Indels 0; Gaps |
| OY | 4 tccccgcagagccccattcccacaagtggcgtggaatgcagaactcggaacccggcgcc 63
 |
| Db | 409 tcgccaggagtgctgttaccacgcagctggcgtggatgatgagctgcagaagccgggggccc 468
 |
| OY | 64 cagatcatatcagtcgtctcatgtcaggaagcacacctgcta 101
 |
| Db | 469 caggtctacatgcattcatgtcaggaagcacacctgcta 506
 |
| RESULT | 7 |
| ID | AAD03321 standard; DNA: 2022 BP. |
| XX | |
| AC | AAD03321; |
| XX | |
| DT | 13-JUN-2001 (first entry) |
| XX | |
| DE | Sus scrofa PRKAG3 splice variant DNA. |
| XX | |
| P15: | gamma subunit; adenosine monophosphate-activated kinase; AMPK; |
| KW | PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic; |
| KW | genetic testing; carbohydrate metabolism disorder; skeletal muscle; |
| KW | cystathione beta synthase; CBS; cardiant; gene therapy; ds. |
| XX | |
| OS | Sus scrofa. |
| XX | |
| Key | Location/Qualifiers |
| TH | |
| CD5 | 1..1545 |
| TT | |

```

FT      /*tag= a  

XX      /product= "Sus scrofa Pkrag3 splice variant"  

PN      WO200120003-A2.  

PD      22-MAR-2001.  

XX      11-SEP-2000; 2000WO-EP09896.  

PF      10-SEP-1999; 99EP-0402236.  

PR      18-MAY-2000; 2000EP-0401388.  

XX      (INRG ) INRA INST NAT RECH AGRONOMIQUE.  

PA      (ANDE/) ANDERSSON L.  

PA      (LOOF/) LOOFT C.  

PA      (KALM/) KALM E.  

XX      Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Galliard C;  

P1      Iannuccielli N, Gellin J, Le Roy P, Chardon P;  

DR      WPI: 2001-244810/25.  

DR      P-PsDB: MAE00224.  

XX      New variants of the gamma subunit of vertebrate adenosine  

PT      monophosphate-activated kinase for diagnosis or treatment of disorders  

PT      associated with energy metabolism such as diabetes, obesity, and  

PT      myopathy -  

PS      Claim 12; Page 69; 71pp: English.  

XX      The present sequence is pig adenosine monophosphate (AMP)-activated  

CC      kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice  

CC      variant DNA. Pkrag3 gene is located in the RN locus of chromosome 15.  

CC      Mutation in Pkrag3 results in an altered regulation of carbohydrate  

CC      metabolism, particularly in skeletal muscle. PRKAG3 is useful as  

CC      therapeutic for treating carbohydrate metabolism disorders such as  

CC      diabetes, obesity, and disorders associated with muscle metabolism  

CC      such as myopathy and cardiovascular diseases, to modulate AMPK  

CC      activity, and for restoring a normal AMPK function. PRKAG3 sequence  

CC      and its functionally altered mutants are useful for the diagnostic  

CC      evaluation, genetic testing and prognosis of a metabolic disorder,  

CC      preferably a carbohydrate metabolism disorder. Primers that can detect  

CC      a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  

CC      useful for detecting a dysfunction of carbohydrate metabolism resulting  

CC      from the expression of a functionally altered allele of PRKAG3.  

CC      Transgenic animal and host cell transformed with PRKAG3 or a  

CC      heterologous AMPK consisting of PRKAG3 or its mutant, are useful for  

CC      screening compounds able to modulate AMPK activity. Nucleic acid  

CC      encoding PRKAG3 is useful for detecting mutations in a Pkrag3 gene, or  

CC      in a sequence encoding the first cystathione beta synthase (CBS) domain  

CC      of PRKAG3 and is useful in gene therapy.  

XX      Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other:  

XX      Query Match          71.7%; Score 72.4; DB 22; Length 2022;  

XX      Best Local Similarity 83.7%; Pred.No. 1.3e-13;  

XX      Matches      82; Conservative    0; Mismatches   16; Indels     0; Gaps    0;  

QY      4 tccccgcaggcccccattcccaactgtggcgtggaatcgacgaacccggcgcc 63  

Db      559 tccccagagtgctgttaccgccagctgttcgtggagtatgatgacctgcagaagccgggggcc 618  

QY      64 cagatcatcacgtcgccttcatacgaggaacacacctgcta 101  

Db      619 caggtctacatgacacttcatacgaggaacacacctgcta 656  

RESULT      8  

AA551418 ID AA551418 standard; DNA: 1428 BP.  

XX      AA551418;  


```

XX 13-FEB-2002 (first entry)
DT
XX
XX Pseudomonas aeruginosa DNA for cellular proliferation protein #3.
DE
XX
XX Antisense: ds: prokaryotic cellular proliferation gene;
KM antibiotic; antibacterial; drug design.
XX
XX Pseudomonas aeruginosa.
OS
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001: 2001WO-US09180.
PF
XX
XX 21-MAR-2000: 2000US-191078P.
PR 23-MAY-2000: 2000US-206848P.
PR 26-MAY-2000: 2000US-207727P.
PR 23-OCT-2000: 2000US-242578P.
PR 27-NOV-2000: 2000US-253625P.
PR 22-DEC-2000: 2000US-257931P.
PR 16-FEB-2001: 2001US-269308P.
XX
XX (ELITR-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH:
XX
XX WPI: 2001-611495/70.
DR P-PSDB: AAU33559.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27: Seq ID No 4000: 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1428 BP: 206 A: 525 C: 423 G: 274 T: 0 other;
SQ

Query Match 29.9%; Score 30.2; DB 23; Length 1428;
Best Local Similarity 58.2%; Pred. No. 3.2;
Matches 53; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 3 gtcccgagcccaattcccaagctggctggtatgacgaatcggaacccggcgc 62
DB 1218 gtccgtatggtgcatatacctgctgacccggcggtgacccgttcacgtgcgcgc 1277
QY 63 ccagatctacatgcgttcacatgacgagc 93
DB 1278 gctgagacatcatgctgacatgacgagc 1308

RESULT 9
AAS74198
ID AAS74198 standard; cDNA: 3918 BP.
XX
XX AAS74198:
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #10002.
DE
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001: 2001WO-US08631.
PF
XX
XX 31-MAR-2000: 2000US-0540217.
PR 23-AUG-2000: 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT:
PI
XX
XX WPI: 2001-639362/73.
DR P-PSDB: ABG10011.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1: SEQ ID No 10002: 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3918 BP: 1328 A: 872 C: 822 G: 896 T: 0 other;
SQ

Query Match 29.1%; Score 29.4; DB 23; Length 3918;
Best Local Similarity 56.8%; Pred. No. 6.9;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 4 tcccgagcccaattcccaagctggctggtatgacgaatcggaacccggcgc 63
DB 2710 tcccttatcccatctcctctctgagtgctggcgacgagcgcggtcgcg 2769
QY 64 cagatctacatgcgttcacatgacgagcactg 98

Db 2770 cagctccacctgcagcccgagctgcgagatccactg 2804

RESULT 10

AA580424
ID AAS80424 standard; cDNA; 3918 BP.

AC AAS80424;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16228.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN W0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG16237.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

Claim 1: SEQ ID No 16228; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques for restoring normal activity of (II) or to treat disease states involving quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations in responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 3918 BP: 1328 A; 872 C; 822 G; 896 T; 0 other;

Query Match 29.1%; Score 29.4; DB 23; Length 3918;
Best Local Similarity 56.8%; Pred. No. 6.9;

Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 4 tccccgagcccatcctccaagctggctggagtcgaactgcggaacccgcgcgc 63
Db 2770 tccctatcccccacatccctcctgctgaggtgtgtggcgcgacagcgagactgcgcgg 2769

OY 64 cagatctacatgcgtctcatgcagagacacactg 98
Db 2770 cagctccacctgcagcccgagctgcgagatccactg 2804

RESULT 11

AA574446
ID AAS74446 standard; cDNA; 4397 BP.

AC AAS74446;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #10250.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN W0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG10259.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

Claim 1: SEQ ID No 10250; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques for restoring normal activity of (II) or to treat disease states involving quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations in responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 4397 BP: 1506 A; 980 C; 935 G; 976 T; 0 other;

Query Match 29.1%; Score 29.4; DB 23; Length 4397;
Best Local Similarity 56.8%; Pred. No. 7;

Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;


```

ABLI9504
ID ABL19504 standard; DNA: 3406 BP.
XX
AC ABL19504:
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9985.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001: 2001WO-US09231.
XX
PR 23-MAR-2000: 2000US-191637P.
XX PK 11-JUL-2000: 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
PI Venter JC, Adams M, Li PMD, Myers EW:
P1 WPI: 2001-6556860/75.
XX
DK
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions .
XX
XX
PS Claim 1: SEQ ID NO 9985; 21pp + Sequence Listing: English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB161840-AB16175) and the encoded proteins
CC (AB157737-AB172072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX
SQ Sequence 3406 BP: 902 A; 852 C; 828 G; 824 T; 0 other;

Query Match 28.7%; Score 29; DB 23; Length 3406;
Best local Similarity 71.7%; Pred. No. 9;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 49 cggaaccgcggccacagctacatgcttcacagcagagacacactgcta 101
   ||| ||||| ||| ||| ||||| ||| ||||| ||| |||
DB 1264 ctgaagccgcggccgcgcgtgtgcgcgcgtcttcaggaagctctccgacta 1316

RESULT 15
ABLI9502
ID ABL19502 standard; DNA: 3593 BP.
XX
AC ABL19502:
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9979.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.

```

| | | |
|----|----|---|
| XX | PN | WO2001/1042-A2. |
| XX | PD | 27-SEP-2001. |
| XX | PF | 23-MAR-2001; 2001WO-US09231. |
| XX | PR | 23-MAR-2000; 2000US-191637P. |
| XX | PR | 11-JUL-2000; 2000US-0614150. |
| PA | | (PEKE) PE CORP NY. |
| XX | PI | Venter JC, Adams M, Li PWD, Myers EW; |
| XX | PI | WPI; 2001-656860/75. |
| XX | DR | |
| XX | PT | New isolated nucleic acid detection reagent for detecting 1000 or more |
| XX | PT | genes from Drosophila and for elucidating cell signalling and cell-cell |
| XX | PT | interactions - |
| PS | | Claim 1; SEQ ID NO 9979; 21pp + Sequence Listing; English. |
| XX | | |
| XX | | The invention relates to an isolated nucleic acid detection reagent |
| CC | CC | capable of detecting 1000 or more genes from Drosophila. The invention is |
| CC | CC | useful in developmental biology and in elucidating cell signalling and |
| CC | CC | cell-cell interactions in higher eukaryotes for the development of |
| CC | CC | insecticides, therapeutics and pharmaceutical drugs. The invention |
| CC | CC | discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA |
| CC | CC | sequences (ABL01840-ABL16175) and the encoded proteins |
| CC | CC | (ABB57737-ABB72072). |
| CC | CC | The sequence data for this patent did not form part of the printed |
| CC | CC | specification, but was obtained in electronic format directly from WIPO |
| XX | XX | at ftp.wipo.int/pub/published_pt_sequences. |
| SO | | Sequence 3593 BP; 920 A; 953 C; 902 G; 818 T; 0 other; |
| | | |
| QY | | Query Match 28.7%; Score 29; DB 23; Length 3593; |
| | | Best Local Similarity 71.7%; Pred.No. 9.1; |
| | | Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0 |
| DG | | 49 cgaagaccgcgcgcagatctactatgcgttcatagcagagacacacctgcta 101 |
| | | |
| DG | | 3377 ctgaagccgcgcgcgcgcgtctgcgcgcgcgtcttcagagactctcagacta 3429 |

Search completed: October 3, 2002, 16:30:46
Job time: 14316 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:04 ; Search time 5701.1 Seconds

(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600

Sequence: 1 cgtgcccgagggccatc.....atgcagagacacccgccta 101

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estow:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 72.2 | 71.5 | 444 | 10 | BF890374 291826 MA |
| 2 | 72.2 | 71.5 | 548 | 10 | BI775360 467815 MA |
| 3 | 67.6 | 66.9 | 685 | 9 | BB630381 BB630381 |
| 4 | 31.4 | 31.1 | 326 | 9 | AI106144 |
| 5 | 30.6 | 30.3 | 679 | 10 | BE540087 |
| 6 | 30.4 | 30.1 | 699 | 10 | BI199916 |
| 7 | 30.2 | 29.9 | 580 | 12 | BH140228 |
| 8 | 30.2 | 29.9 | 894 | 12 | CNS01MA1 |
| 9 | 29.8 | 29.5 | 910 | 12 | CNS02SYT |
| 10 | 29.8 | 29.5 | 1010 | 12 | CNS03856 |
| 11 | 29.6 | 29.3 | 655 | 10 | BM133818 |
| 12 | 29.6 | 29.3 | 930 | 10 | BE269692 |
| 13 | 29.4 | 29.1 | 955 | 10 | BE269692 |
| 14 | 29.2 | 28.9 | 1000 | 10 | BM468546 |
| 15 | 29.2 | 28.7 | 313 | 9 | AA264931 |
| 16 | 29.2 | 28.7 | 436 | 10 | T95291 |
| 17 | 29.2 | 28.7 | 546 | 9 | AA246471 |

| | | | | | |
|----|------|------|-----|----|----------|
| 18 | 29 | 28.7 | 554 | 9 | A1260567 |
| 19 | 29 | 28.7 | 561 | 10 | B1229683 |
| 20 | 29 | 28.7 | 583 | 10 | B1232284 |
| 21 | 29 | 28.7 | 589 | 9 | AA539970 |
| 22 | 29 | 28.7 | 627 | 10 | B1633071 |
| 23 | 29 | 28.7 | 629 | 9 | A1259247 |
| 24 | 29 | 28.7 | 666 | 10 | B1242279 |
| 25 | 29 | 28.7 | 738 | 9 | A1544103 |
| 26 | 28.8 | 28.5 | 569 | 10 | BM14376 |
| 27 | 28.8 | 28.5 | 750 | 10 | BM019844 |
| 28 | 28.6 | 28.3 | 339 | 10 | BF773519 |
| 29 | 28.6 | 28.3 | 445 | 10 | BG558441 |
| 30 | 28.6 | 28.3 | 448 | 9 | AM497449 |
| 31 | 28.6 | 28.3 | 509 | 10 | BG560394 |
| 32 | 28.6 | 28.3 | 591 | 9 | AM678584 |
| 33 | 28.6 | 28.3 | 593 | 12 | BH140663 |
| 34 | 28.6 | 28.3 | 605 | 9 | AA195150 |
| 35 | 28.6 | 28.3 | 707 | 12 | AG130442 |
| 36 | 28.6 | 28.3 | 718 | 9 | AV753229 |
| 37 | 28.6 | 28.3 | 742 | 10 | BE394076 |
| 38 | 28.6 | 28.3 | 904 | 10 | BF033403 |
| 39 | 28.4 | 28.1 | 260 | 9 | AM816448 |
| 40 | 28.4 | 28.1 | 712 | 9 | BB623242 |
| 41 | 28.4 | 28.1 | 812 | 9 | AA519383 |
| 42 | 28.2 | 27.9 | 517 | 9 | AM631272 |
| 43 | 28.2 | 27.9 | 525 | 9 | AA922163 |
| 44 | 28.2 | 27.9 | 592 | 12 | AZ870322 |
| 45 | 28.2 | 27.9 | 802 | 12 | BH077939 |

ALIGNMENTS

RESULT 1
LOCUS BF890374 444 bp mRNA linear EST 25-APR-2001
DEFINITION 291826 MARC 3BOV Bos taurus CDNA 5', mRNA sequence.
ACCESSION BF890374
VERSION BF890374.1 GI:12281760
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Peters,G., Holt,I., Karaycheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.

TITLE
JOURNAL
MEDLINE
COMMENT Sequence evaluation of four pooled-tissue normalized bovine CDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_c trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.

PCR Primers

FORWARD: AGCAACACGTATGACAT
BACKWARD: GTTTCACGTACGACG
Plate: 57 row: H column: 11
Seq primer: ATTACGTACTACTATAG.

FEATURES

Location/Qualifiers
1..444
/organism="Bos taurus"
/db_xref="taxon:9913"

| FEATURES | source |
|--------------------------|---|
| Query Match | 66.9%: Score 67.6; DB 9; Length 685; |
| Best Local Similarity | 80.6%: Pred. No. 1.7e-09; |
| Matches 79: Conservative | 0: Mismatches 19; Indels 0; Gaps 0; |
| Yy | 4 tcccccaagccccatllcccaagctgagctgagtgagcaactgcggaaacccggccg 63 |
| Db | 496 tcccccaagcccccaacttcttgccctggcattggcattgacacacttcacaaacccggagcc 555 |
| Oy | 64 cagatctaaatgcgtctcatgacgaggaacacccctccta 101 |
| Db | 556 caggtctacatgaccttattgacgagacacacctgtta 593 |
| RESULT | 4 |
| LOCUS | A1106144/c |
| DEFINITION | A1106144 326 bp mRNA linear EST 25-AUG-1998 |
| ACCESSION | db03dl2_p1 2F adult heart library Danio rerio cDNA 5 prime similar |
| VERSION | A1106144 |
| KEYWORDS | A1106144.1 GI:3461247 |
| SOURCE | EST. |
| ORGANISM | zebrafish. |
| REFERENCE | Danio rerio. |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. |
| TITLE | 1 (bases 1 to 326). |
| JOURNAL | Chen, J.N., Desauvage, F., Hosobuchi, M., Jackson, D.G. and Fishman |
| COMMENT | M.C. Expressed Sequences from The Adult Zebrafish Heart unpublished (1998) |
| | Contact: Mark C. Fishman |
| | Cardiovascular Research Center |
| | Massachusetts General Hospital |
| | Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA |
| | Fax: 617/265806 |
| | Email: fishman@gh.cvr.harvard.edu |
| | http://zebrafish.mgh.harvard.edu |
| | The original clones used for sequencing are no longer available; |
| | the library is available from Mark C. Fishman. |
| | Insert length: 326 Std Error: 0.00 |
| | Seq primer: pl. |
| | Location:Qualifiers |
| | 1..326 |

| | | | | | |
|-----------------------|---|---|--------------|------------|-------------------|
| BASE COUNT | 83 a | 66 c | 96 g | 79 t | 2 others |
| ORIGIN | | | | | |
| Query Match | 31.1%: Score 31.4; DB 9; Length 326; | | | | |
| Best Local Similarity | 62.7%: Pred No. 34; | | | | |
| Matches | 47; Conservative | 0; Mismatches | 26; Indels | 0; Gaps | 0; |
| Qy | 23 | ccaagctggagctggagctgcgaactgcggaacccgcgcgccagatctacatcgcttca | 82 | | |
| Db | 256 | CCAAACTGAAGGCTCTTGTGCTTCNCAGACACTGCGGACGATCAACAGATTGCTCCA | 197 | | |
| Qy | 83 | tcgagagacacacct | 97 | | |
| Db | 196 | GGCAGCTTCACACT | 182 | | |
| RESULT | 5 | BE540087 | 679 bp | MRNA | linear |
| LOCUS | BE540087/c | 601060879P2 NIH_MGC_10 | Homo sapiens | CDNA clone | IMAGE:3447378 5', |
| DEFINITION | mRNA sequence. | | | | |
| ACCESSION | BE540087 | BE540087.1 | GI:9768732 | | |
| VERSION | EST. | | | | |
| KEYWORDS | human. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| TITLE | NIH-MGC http://mgc.nci.nih.gov/ . | | | | |
| JOURNAL | National Institutes of Health, Mammalian Gene Collection (MGC) | | | | |
| COMMENT | Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LRAM8420 row: n column: 19
High quality sequence stop: 668.
Location/Qualifiers
1. 679
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3447378"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH108"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies." | | | | |
| BASE COUNT | 124 a | 226 c | 177 g | 152 t | |
| ORIGIN | | | | | |
| Query Match | 30.3%: Score 30.6; DB 10; Length 679; | | | | |
| Best Local Similarity | 62.3%: Pred. No. 68; | | | | |
| Matches | 48; Conservative | 0; Mismatches | 29; Indels | 0; Gaps | 0; |

```

OY      9  gcaagcccaatcccaagctggagatgacgaactgcgaaccggcgccacgat 68
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      244 GCAACCCGCGCTTCTGCGAGGTGTCAGAGACGCAACTGGGACAGCTGGCGCTGCAT 185
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      69  ctacatgcgttcatagc 85
        || | | |||||
Db      184 TCACTGCACATACATGC 168
        || | | |||||

RESULT  6
B1199916/c 699 bp mRNA linear EST 10-JUL-2001
LOCUS      602760950F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4896324 5',
DEFINITION mRNA sequence.
ACCESSION B1199916
VERSION   B1199916.1 GI:14654937
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabsr@mail.nih.gov
           Tissue Procurement: ATCC
           CDNA Library Preparation: Ling Hong/Rubin Laboratory
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
           Plate: LLCMT783 row: c column: 13
           High quality sequence start: 5
           High quality sequence stop: 669.
FEATURES
  Source
    1..699
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4896324"
    /clone_1lb="NIH_MGC_19"
    /tissue_type="neuroblastoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: Brain; Vector: pORF7; Site 1: XhoI; Site 2:
    EcoRI; cDNA made by oligo-dT priming. Directionally
    cloned into EcoRI/XhoI sites using the following 5'
    adaptor: GGCACGAG(C). Library constructed by Ling Hong
    in the laboratory of Gerald M. Rubin (University of
    California, Berkeley) using ZAP-cDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies).
    Note: this is a NIH_MGC Library."
BASE COUNT 120 a 216 c 222 g 141 t
ORIGIN
Query Match 30.1%; Score 30.4; DB 10; Length 699;
Best Local Similarity 67.2%; Pred. No. 78;
Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
OY      36  ggtatgcgaactcgcgaaccggcgccagatctacatgcgttcatagcagagacac 95
        ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      78  GGACGACGACTACGAGATCCGAGATCCGGAACCTCCGGAGCGGATGCACGCGCTCAC 19
        ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      96  ctgc 99
        || |
Db      18  CTCG 15

RESULT  7
B1140228/c 580 bp DNA linear GSS 07-AUG-2001
LOCUS      B1140228

```

```

DEFINITION ZM8BB0001118f Maize B73 Zea mays genomic clone ZM8BB0001118f, DNA
sequence.
ACCESSION B1140228
VERSION   B1140228.1 GI:15099289
KEYWORDS  GSS.
SOURCE    Zea mays.
ORGANISM  Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE Tomkins,J.P., Main,D., Golcochea,J.L., Frisch,D.A. and Wing,R.A.
AUTHORS   A Deep-Coverage BAC Library for Maize
JOURNAL   Unpublished (2001)
COMMENT    Contact: Wing RA
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Tel: 864 656 7288
           Fax: 864 656 4293
           Email: twing@clemson.edu
           Seq primer: TAATGACTCCTATGCGC
           Class: BAC ends
           High quality sequence start: 71
           High quality sequence stop: 510.
FEATURES
  source
    1..580
    /organism="Zea mays"
    /strain="B73"
    /cultivar="B73"
    /db_xref="taxon:4577"
    /clone="ZM8BB0001118f"
    /clone_1lb="Maize B73"
    /tissue_type="Young leaves"
    /lab_host="E. coli"
    /note="Vector: pCUGIBAC-1; Site 1: HindIII; Site 2: NotI;
    For more details on library preparation, ordering clones
    and sequence analysis see
    http://www.genome.clemson.edu/projects/stc/maize/ZM8BB"
BASE COUNT 74 a 221 c 138 g 146 t
ORIGIN
Query Match 29.9%; Score 30.2; DB 12; Length 580;
Best Local Similarity 62.7%; Pred. No. 85;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
OY      27  gctgggctggatgacgaactcgcgaaccggcgccagatctacatgcgttcatagc 86
        || | ||||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db      147 GCCGACTCGGACGACAACTGAGGAGCGCGCTCGAGGGGCTCCACCGCAGTAGGAA 88
        ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
OY      87  ggaagcacacctgcta 101
        || | |||||
Db      87  GTTTCGACCGGCGA 73

RESULT  8
CNS01M1/c 894 bp DNA linear GSS 14-JUN-2001
LOCUS      CNS01M1
DEFINITION Anopheles gambiae GSS SP6 end of clone 21E19 of NotreDamel library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION AL150666
VERSION   AL150666.1 GI:7011145
KEYWORDS  GSS.
SOURCE    African malaria mosquito.
ORGANISM  Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
Culicoidae; Anopheles.
REFERENCE 1 (bases 1 to 894)
AUTHORS   Genoscope.
TITLE     Direct Submission

```


| | Query Match | 29.5% | Score 29.8 | DB 12 | length 1010 |
|----|-----------------------|--|-------------------|-------|---------------|
| | Best Local Similarity | 66.2% | Pred. No. 1.3e+02 | | |
| | Matches | 43 | Conservative | 0 | Mismatches 22 |
| | | | | | Indels 0 |
| | | | | | Gaps 0 |
| Oy | 30 | ggggcggagatgacgaactgcgggaacccgcgcggccagatctacatgcgttatgcaga | 89 | | |
| Db | 620 | ggcgatgaaagcgccttccacacgtacccccaagcactctgacattatccgagccatcagaa | 679 | | |
| Oy | 90 | ggcaaa | 94 | | |
| Db | 680 | gcaca | 684 | | |

| RESULT | 11 | | | | |
|------------|--|-------------|---------------------------------|-------------|-----------------|
| LOCUS | BM133818 | | | | |
| DEFINITION | BM133818 | 655 bp | mRNA | linear | EST 07-JAN-2002 |
| LOCUS | NX1V_012_F09.F | NX1V | (Nsif Xylem Late wood vertical) | Pinus taeda | CDNNA |
| ACCESSION | clone NX1V_012_F09 5' | | | | |
| VERSION | BM133818 | | | | |
| KEYWORDS | BM133818.1 | GI:17141057 | | | |
| SOURCE | EST. | | | | |
| ORGANISM | loblolly pine. | | | | |
| | Pinus taeda | | | | |
| REFERENCE | Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 655) | | | | |
| AUTHORS | Sederoff, R. | | | | |
| TITLE | Molecular Basis of Wood Formation in the Pine Megagenome | | | | |
| JOURNAL | Unpublished (2000) | | | | |
| COMMENT | Contact: Johnson, Arthur | | | | |

Seq primer: T3.
Location/Qualifiers
1. .655
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3152"
/clone="NXLV.012.P03"
/clone_id="NXLV (NsI
/tissue_type="primary xylem"
/dev_stage="late wood"
/lab_host="XLI-Blue"
/note="Vector: pRIPlex; Site: 1; EcoRI: The library is from late (summer-August) wood, taken from below the crown of a 20 year old tree. The harvested xylem tissue was on the cusp between transitional and mature wood. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGGCCCATTAATGCC'."

| Query Match | 29.3%; | Score 29.6; | DB 10; | Length 655; |
|-----------------------|--|--|-----------|-----------------|
| Best Local Similarity | 57.6%; | Pred. No. 1.3e+02; | | |
| Matches 53; | Conservative 0; | Mismatches 39; | Indels 0; | Gaps 0; |
| OY | 4 | tcgccgcaggccccattcccaagctggtgcgtgagatgacgaactgcgtggaacccgcgcgc | 63 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 235 | ttcccgggaaacccaattttccagAACCAAGGTAGTATCCCAACCCCGCAGCCCTAATCA | 294 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| OY | 64 | cagatctacatgctgcgtctcatcagtgagagacac | 95 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 295 | CAGTTCATCATGCCCATCTACAATTTCACAC | 326 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| RESULT 12 | | | | |
| LOCUS | BE269692 | 930 bp | mrna | linear |
| DEFINITION | 601185836f1 NIH_MGC_8 Homo sapiens CDNA clone IMAGE:3543584 5' | | | EST 13-JUL-2000 |

| | |
|-----------|--|
| ACCESSION | mRNA sequence. |
| VERSION | BE269692 |
| KEYWORDS | BE269692.1 GI:9143317 |
| SOURCE | EST. |
| ORGANISM | human. |
| REFERENCE | Homo sapiens |
| TITLE | Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi, |
| AUTHORS | Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo. |
| JOURNAL | 1 (bases 1 to 930) |
| COMMENT | NIH-MGC http://mgc.nci.nih.gov/ .
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: L1CM237 row: 9 column: 09
High quality sequence, stop: 238. |

| FEATURES | SOURCE |
|--|-------------------|
| Location/Qualifiers | |
| 1..930 | |
| /organism="Homo sapiens" | |
| /db_xref="taxon:9606" | |
| /clone="IMAGE:3543584" | |
| /clone_id="NH_MGC_8" | |
| /tissue_type="Burkitt lymphoma" | |
| /lab_host="DH10B (phage-resistant)" | |
| /note="Organ: lymph; Vector: pORF7; Site:1: Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." | |
| BASE COUNT | |
| ORIGIN | |
| 226 a | 228 c 267 g 209 t |

| RESULT | 13 |
|------------|---|
| LOCUS | BC394974 |
| DEFINITION | BC394974 955 bp mRNA linear EST 12-MAR-2002 |
| ACCESSION | 602427369F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4578696 5', mRNA sequence. |
| VERSION | BC394974 |
| KEYWORDS | BC394974.1 GI:13288422 |
| SOURCE | EST. |
| ORGANISM | human. |
| REFERENCE | Homo sapiens |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| TITLE | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| JOURNAL | 1 (bases 1 to 955) |
| COMMENT | NIH-MGC http://mgc.nci.nih.gov/ .
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov |

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LNCM1297 row: 3 column: 17
High quality sequence stop: 717.
Location/Qualifiers
1. .955

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4579696"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAC(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 268 a 208 c 225 g 254 t
ORIGIN

Query Match 29.1% Score 29.4; DB 10; Length 955;
Best Local Similarity 66.7% Pred. No. 1.6e+02;
Matches 42; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 33 ctggatgacgaactgcggaaaccggccagatctacatgcgtcattcagagaca 92
|||
Db 836 CTGAGTACGCTACTCCGAAATCTGCGCGCGAAMAAACCGTTTCATGAAGTTGA 895
QY 93 cac 95
|||
Db 896 CAC 898

RESULT 14
BM468546 1000 bp mRNA linear EST 05-FEB-2002
LOCUS
DEFINITION
AGENCOURT_6475521 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578160
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
BM468546.1 GI:18517588
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1000)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCDB/DP
CDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM12332 row: m column: 09
High quality sequence stop: 725.
Location/Qualifiers
1. .1000

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578160"

/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT 276 a 244 c 303 g 174 t 3 others
ORIGIN

Query Match 28.9% Score 29.2; DB 10; Length 1000;
Best Local Similarity 64.2% Pred. No. 1.9e+02;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 22 ccgaagctggctggatgaactgcggaaaccggccagatctacatgccttc 81
|||
Db 690 CACAGATGCTGTATGAGAGGACCTCCAGAACTCTGTANACAGACCAGATGAGCTCC 749
QY 82 atgcag 88
|||
Db 750 CAGCAGG 756

RESULT 15
AA264931 313 bp mRNA linear EST 19-APR-2001
LOCUS
DEFINITION
LD08544.5prime LD Drosophila melanogaster embryo Bluescript.
Drosophila melanogaster cDNA clone LD08544 5prime, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
AA264931.1 GI:1900974
EST.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 313)
AUTHORS
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMI Drosophila EST Project
TITLE
JOURNAL
Unpublished (2001)
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Blast expect value = 1.1E-117 on X03121: Drosophila melanogaster
serendipity (sry) locus DNA sequence
Plate: 85 row: D column: 8
High quality sequence stop: 240.
Location/Qualifiers
1. .313

FEATURES
source
/organism="Drosophila melanogaster"
/db_xref="BDGP_EST:BDCLN007879"
/db_xref="taxon:7227"
/clone="LD08544"
/clone_lib="LD Drosophila melanogaster embryo Bluescript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site:1: EcoRI; Site:2: XhoI; Constructed using Stratagene ZAP-cDNA synthesis kit. Oligo dT-primed and directionally cloned at EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT 72 a 99 c 82 g 60 t
ORIGIN

Query Match 28.7% Score 29; DB 9; Length 313;
Best Local Similarity 71.7% Pred. No. 1.6e+02;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 49 cygaanaccgcgcgccagatctacatgcgcttcacatgcagagcacacctgcta 101
1 111 111111111 1 1 1 1111 1 1111111 1 11111
Db 235 CTGAAGCCCGCGCGCGCTGTGCCCGCGCTTTCAGGAGCTCTCCGACTA 287

Search completed: October 3, 2002, 16:15:08
Job time: 16963 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:22:14 : Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600

Perfect score: 101
Sequence: 1 cgtccccgcagagcccaatt.....atgcagagacacactgtcta 101

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767056

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCrUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|---------------------|--------------------|
| 1 | 26.4 | 26.1 | 44377 | 2 US-08-804-227C-7 | Sequence 7, Appl1 |
| 2 | 26.4 | 26.1 | 44377 | 2 US-08-804-198-1 | Sequence 1, Appl1 |
| 3 | 26 | 25.7 | 4403765 | 4 US-09-103-840A-2 | Sequence 2, Appl1 |
| 4 | 25.8 | 25.5 | 594 | 4 US-08-441-507-17 | Sequence 17, Appl1 |
| 5 | 25.8 | 25.5 | 6085 | 4 US-09-029-603-4 | Sequence 4, Appl1 |
| 6 | 25.2 | 25.0 | 2980 | 2 US-08-463-081B-13 | Sequence 13, Appl1 |
| 7 | 25.2 | 25.0 | 2980 | 2 US-08-461-379A-13 | Sequence 13, Appl1 |
| 8 | 25.2 | 25.0 | 2980 | 2 US-08-462-390B-13 | Sequence 13, Appl1 |
| 9 | 25.2 | 25.0 | 2980 | 3 US-08-463-074B-13 | Sequence 13, Appl1 |
| 10 | 25.2 | 25.0 | 2980 | 3 US-08-465-585C-13 | Sequence 13, Appl1 |
| 11 | 25.2 | 25.0 | 2980 | 3 US-08-652-446-13 | Sequence 13, Appl1 |
| 12 | 24.8 | 24.6 | 255 | 4 US-08-980-823-31 | Sequence 31, Appl1 |
| 13 | 24.8 | 24.6 | 1364 | 1 US-08-306-691B-50 | Sequence 50, Appl1 |
| 14 | 24.8 | 24.6 | 1364 | 5 PCT-US93-06251-65 | Sequence 65, Appl1 |
| 15 | 24.8 | 24.6 | 1900 | 4 US-09-608-285A-47 | Sequence 47, Appl1 |
| 16 | 24.8 | 24.6 | 2294 | 4 US-09-608-285A-49 | Sequence 49, Appl1 |
| 17 | 24.8 | 24.6 | 2371 | 4 US-09-608-285A-46 | Sequence 46, Appl1 |
| 18 | 24.8 | 24.6 | 2497 | 4 US-09-608-285A-51 | Sequence 51, Appl1 |
| 19 | 24.8 | 24.6 | 2693 | 4 US-09-608-285A-48 | Sequence 48, Appl1 |
| 20 | 24.8 | 24.6 | 2721 | 6 5215881-2 | Patent No. 5215881 |
| 21 | 24.8 | 24.6 | 2762 | 4 US-09-608-285A-26 | Sequence 26, Appl1 |
| 22 | 24.8 | 24.6 | 2762 | 4 US-09-608-285A-52 | Sequence 52, Appl1 |
| 23 | 24.8 | 24.6 | 2762 | 4 US-09-240-639-1 | Sequence 1, Appl1 |
| 24 | 24.8 | 24.6 | 3123 | 1 US-08-094-889-2 | Sequence 2, Appl1 |
| 25 | 24.8 | 24.6 | 8438 | 1 US-07-945-283-1 | Sequence 1, Appl1 |
| 26 | 24.6 | 24.4 | 30001 | 1 US-08-125-468-1 | Sequence 1, Appl1 |
| 27 | 24.6 | 24.4 | 30001 | 2 US-08-474-933-1 | Sequence 1, Appl1 |

ALIGNMENTS

| | | | | | |
|----|------|------|-------|---------------------|--------------------|
| 28 | 24.4 | 24.2 | 2327 | 4 US-08-868-435-1 | Sequence 1, Appl1 |
| 29 | 24.4 | 24.2 | 2327 | 4 US-08-744-231-1 | Sequence 1, Appl1 |
| 30 | 24.4 | 24.2 | 31571 | 1 US-08-323-443B-1 | Sequence 1, Appl1 |
| 31 | 24.4 | 24.2 | 68750 | 3 US-09-335-409-1 | Sequence 1, Appl1 |
| 32 | 24.4 | 24.2 | 68750 | 4 US-09-568-102-1 | Sequence 1, Appl1 |
| 33 | 24.4 | 24.2 | 68750 | 4 US-09-567-969-1 | Sequence 1, Appl1 |
| 34 | 24.4 | 24.2 | 68750 | 4 US-09-568-480-1 | Sequence 1, Appl1 |
| 35 | 24.4 | 24.2 | 68750 | 4 US-09-568-486-1 | Sequence 1, Appl1 |
| 36 | 24.4 | 24.2 | 68750 | 4 US-09-568-472-1 | Sequence 1, Appl1 |
| 37 | 24.4 | 24.2 | 71989 | 4 US-09-443-501A-2 | Sequence 2, Appl1 |
| 38 | 24.2 | 24.0 | 4695 | 2 US-08-231-193A-57 | Sequence 57, Appl1 |
| 39 | 24.2 | 24.0 | 4695 | 2 US-08-486-273A-57 | Sequence 57, Appl1 |
| 40 | 24.2 | 24.0 | 4695 | 3 US-08-940-086A-57 | Sequence 57, Appl1 |
| 41 | 24.2 | 24.0 | 4695 | 4 US-08-940-035A-57 | Sequence 57, Appl1 |
| 42 | 24.2 | 23.8 | 662 | 4 US-08-441-507-1 | Sequence 1, Appl1 |
| 43 | 23.8 | 23.6 | 1406 | 4 US-09-287-097-1 | Sequence 1, Appl1 |
| 44 | 23.8 | 23.6 | 3048 | 5 PCT-US95-14418-1 | Sequence 1, Appl1 |
| 45 | 23.8 | 23.6 | 3048 | 5 PCT-US95-15327-1 | Sequence 1, Appl1 |

RESULT 1

US-08-804-227C-7
Sequence 7, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuestoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: Linear
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:

```

: NAME/KEY: CDS
: LOCATION: 31329..36071
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 36155..41830
: US-08-804-227C-7
:
Query Match      26.1%; Score 26.4; DB 2; Length 44377;
Best Local Similarity 69.2%; Pred. No. 17;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 22 ccgaagctggctggtgacgaactgcggaaccgcgcgcacatctaca 73
DB 27484 CCGCTGCTGGGAGCGGCGTCAACTGCCGAGTCCGGTGACCGGATGTACA 27535

RESULT 2
US-08-804-198-1
: Sequence 1, Application US/08804198
: Patent No. 5945320
: GENERAL INFORMATION:
: APPLICANT: Burgett, Stanley G.
: APPLICANT: Kuhstoss, Stuart A.
: APPLICANT: Rao, Nagaraja R.
: APPLICANT: Richardson, Mark A.
: APPLICANT: Rostock, Paul R., Jr.
: TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PAUL R. CANTRELL, 1138
: STREET: LILLY CORPORATE CENTER
: CITY: INDIANAPOLIS
: STATE: IN
: COUNTRY: USA
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Macintosh
: SOFTWARE: Microsoft Word 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/804,198
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: CANTRELL, PAUL R.
: REGISTRATION NUMBER: 36,470
: REFERENCE/DOCKET NUMBER: P9113
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3885
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 44377 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 350..14002
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 14046..20036
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 20110..31284
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31329..36071
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 36155..41830
```

```

US-08-804-198-1

Query Match      26.1%; Score 26.4; DB 2; Length 44377;
Best Local Similarity 69.2%; Pred. No. 17;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 22 ccgaagctggctggtgacgaactgcggaaccgcgcgcacatctaca 73
DB 27484 CCGCTGCTGGGAGCGGCGTCAACTGCCGAGTCCGGTGACCGGATGTACA 27535

RESULT 3
US-09-103-840A-2
: Sequence 2, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO: 2
: LENGTH: 4403765
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      25.7%; Score 26; DB 4; Length 4403765;
Best Local Similarity 58.8%; Pred. No. 40;
Matches 47; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 6 ccgcgaagcccaattcccaagctggctggatgcgaactgcggaaccgcgcacca 65
DB 2113570 ccgcgaagcgtggtatcgcgaaccgcgttacgctgacgagatctgcagccgcgcacca 2113629

OY 66 gatcatatggcgttcattgc 85
DB 2113630 gaaccggaatggtcagctgcgc 2113649

RESULT 4
US-08-441-507-17
: Sequence 17, Application US/08441507
: Patent No. 6214358
: GENERAL INFORMATION:
: APPLICANT: Singh, Mohan Bir.
: APPLICANT: Smith, Penelope; and
: APPLICANT: Knox, Robert Bruce
: TITLE OF INVENTION: Protein Allergens of the Species Cynodon
: TITLE OF INVENTION: Dactylon
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
```

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/441,507
: FILING DATE: 15-May-1995
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 07/969,875
: FILING DATE: 30-October-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragouras, Amy E.
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: IMI-049DV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 742-4214
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 594 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-441-507-17
```

```

Query Match      25.5%; Score 25.8; DB 4; Length 594;
Best Local Similarity 63.9%; Pred. No. 9,7;
Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```

QY      8 cgcagagcccaattcccaagctggctggatgacgaactgcggaaccgcgcacaga 67
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      17 CGTTCCGCCCATGCGCCAGAGAGCGGAGAGACCAACTGCGCAGCGGAGACTGA 76
QY      68 t 68
DB      77 T 77
```

```

RESULT      5
US-09-029-603-4
: Sequence 4, Application US/09029603
: Patent No. 6210935
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Engel, Natalie
: APPLICANT: Bietenhader, Jurg
: APPLICANT: Toupet, Christine
: APPLICANT: Pospiech, Andreas
: TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
: FILE REFERENCE: 4-20555/A/PCT
: CURRENT APPLICATION NUMBER: US/09/029,603
: CURRENT FILING DATE: 1998-03-20
: EARLIER APPLICATION NUMBER: PCT/EP96/03643
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 6085
: TYPE: DNA
: ORGANISM: Streptomyces longisporoflavus
: FEATURE:
: NAME/KEY: misc_RNA
: LOCATION: (378)..(1665)
: OTHER INFORMATION: ORF
: FEATURE:
: NAME/KEY: misc_RNA
: LOCATION: (1747)..(2553)
: OTHER INFORMATION: ORF
: NAME/KEY: misc_RNA
: LOCATION: (2593)..(4011)
: OTHER INFORMATION: ORF
```

```

: FEATURE:
: NAME/KEY: misc_RNA
: LOCATION: (4013)..(4999)
: OTHER INFORMATION: ORF
: FEATURE:
: NAME/KEY: misc_RNA
: LOCATION: (5071)..(6085)
: OTHER INFORMATION: ORF
: US-09-029-603-4
```

```

Query Match      25.5%; Score 25.8; DB 4; Length 6085;
Best Local Similarity 56.5%; Pred. No. 17;
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
```

```

QY      15 cccattcccaagctggctggatgacgaactgcggaaccgcgcacagatctacat 74
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      3243 ccgctgctcaccctcgccagactgcgacgacgagacgacctggtcaacat 3302
QY      75 gcgcctcatgacgagacacactgc 99
      1 | | | | | | | | | | | | | | | | | | | | | |
DB      3303 gaacgcagagacggtgctgctgctgc 3327
```

```

RESULT      6
US-08-463-081B-13/C
: Sequence 13, Application US/08463081B
: Patent No. 5871960
: Patent No. 5871960 5837487
: GENERAL INFORMATION:
: APPLICANT: Smith, Kendall A. & Beadling, Carol
: TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PRETTY, SCHROEDER & POPLANSKI
: STREET: 444 South Flower St. - Suite 1900
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0,
: SOFTWARE: Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/463,081B
: FILING DATE: 5-JUN-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/104,736
: FILING DATE: 10-AUG-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/796,066
: FILING DATE: 20-NOV-91
: ATTORNEY/AGENT INFORMATION:
: NAME: Viviana Amzel, Ph. D.
: REGISTRATION NUMBER: 30,930
: REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 622-7700
: TELEFAX: (213) 489-4210
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2980 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 240..1475
```

US-08-463-081B-13

Query Match 25.0%: Score 25.2; DB 2; Length 2980;
Best Local Similarity 71.7%: Pred. No. 22;
Matches 33: Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 20 ttcccaagctggctggatgacgaactgagaaacccgagccca 65
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2119 TTCGAGCGAGGGGTGGGAGCAGACTGCGGAACCTCCCCCAA 2074

RESULT 7

US-08-461-379A-13/C
; Sequence 13, Application US/08461379A
; Patent No. 5871961
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; VECTOR AND TRANSFORMED CELL THEREOF, AND
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,379A
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108; 08/104,736
; APPLICATION NUMBER: 6 07/796,066
; FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)470-0700
; TELEFAX: (610)470-0701
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..1475
US-08-461-379A-13

Query Match 25.0%: Score 25.2; DB 2; Length 2980;
Best Local Similarity 71.7%: Pred. No. 22;
Matches 33: Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 20 ttcccaagctggctggatgacgaactgagaaacccgagccca 65
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2119 TTCGAGCGAGGGGTGGGAGCAGACTGCGGAACCTCCCCCAA 2074

RESULT 8

US-08-462-390B-13/C

; Sequence 13, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
; APPLICANT: Smith, K. A. & Beadling, C.
; TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
; TRANSFORMED CELL THEREOF, AND EXPRESSION THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,390B
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108
; FILING DATE: 27-OCT-1994
; APPLICATION NUMBER: USSN 08/104,736
; FILING DATE: 10-AUG-1993
; APPLICATION NUMBER: USSN 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)407-0700
; TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..1475
US-08-462-390B-13

Query Match 25.0%: Score 25.2; DB 2; Length 2980;
Best Local Similarity 71.7%: Pred. No. 22;
Matches 33: Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 20 ttcccaagctggctggatgacgaactgagaaacccgagccca 65
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2119 TTCGAGCGAGGGGTGGGAGCAGACTGCGGAACCTCCCCCAA 2074

RESULT 9

US-08-463-074B-13/C
; Sequence 13, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector and
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
444 South Flower St. - Suite 1


```
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,585
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/463,081
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/461,379
; FILING DATE: 5-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/739,523
; FILING DATE: 29-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: FP66 40035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..1475
; US-08-652-446-13
```

```
Query Match 25.0%; Score 25.2; DB 3; Length 2980;
Best Local Similarity 71.7%; Pred. No. 22;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 20 ttcccaagctggctgagatgacactgcggaaccggcgccca 65
||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2119 TTCGAGCGCAGGGGTGGCGAGCAGACCTCGGAAACCTCCCA 2074
```

```
RESULT 12
US-08-990-823-31
; Sequence 31, Application US/08990823D
; Patent No. 6228371
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; FILE REFERENCE: 49086
; CURRENT APPLICATION NUMBER: US/08/990,823D
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: US 96/10375
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,254
; EARLIER FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: Modified base
; OTHER INFORMATION: n represents a or g or c or t/u
US-08-990-823-31
```

```
Query Match 24.6%; Score 24.8; DB 4; Length 255;
Best Local Similarity 53.0%; Pred. No. 17;
Matches 53; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
OY 2 tgtcccgacgagcccatctcccaagctgggtgatgacgaactgcggaaccggcg 61
||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 143 tgcctcgctgcctgcctgacgaagcgccgcgacatctctgcgcagagcgtggcg 202
OY 62 ccagatctacatgctgcttcacgacgagacacactgcta 101
||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 203 ggcaggtcgacatcttcgatgtcaatgtgcccacatcca 242
```

```
RESULT 13
US-08-306-691B-50/C
; Sequence 50, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skoraki, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-50
```

```
Query Match 24.6%; Score 24.8; DB 1; Length 1364;
Best Local Similarity 54.3%; Pred. No. 25;
Matches 50; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
OY 9 ggaagcccatctcccaagctgggtgatgacgaactgcggaaccggcgcccgat 68
||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 529 GGAGACCTGGTGGCCAAACGTGGGATGGGTGGCCACCATGGGGTATGAGAGCTGCAT 470
OY 69 ctacatgcgtcatgacgagacacactgct 100
||| ||||| ||||| ||||| ||||| ||||| |||||
DB 469 CTGCTCCAGCTCATGTGGCGGTAGAGCTGCT 438
```

```
RESULT 14
PCT-US93-06251-65/C
; Sequence 65, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
```

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
 TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
 STREET: 400 garden City plaza
 CITY: garden City
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/06251
 FILING DATE: 19930630
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: DIGILIO, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8586
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 65:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1364 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PCT-US93-06251-65

RESULT 15
US-09-608-285A-47/C
Sequence 47: Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
TITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16

```

1  PRIOR APPLICATION NUMBER: 09/350,836
2  PRIOR FILING DATE: 1999-07-09
3  PRIOR APPLICATION NUMBER: 09/273,447
4  PRIOR FILING DATE: 1999-03-19
5  PRIOR APPLICATION NUMBER: 09/244,444
6  PRIOR FILING DATE: 1999-02-04
7  PRIOR APPLICATION NUMBER: 09/122,449
8  PRIOR FILING DATE: 1998-07-24
9  PRIOR APPLICATION NUMBER: 09/118,205
10 PRIOR FILING DATE: 1998-07-16
11 NUMBER OF SEQ ID NOS: 60
12 SOFTWARE: PatentIn Ver. 2.0
13 SEQ ID NO 47
14 LENGTH: 1900
15 TYPE: DNA
16 ORGANISM: Homo sapiens
17 IS-09-608-2854-47

```

| Query Match | 24.6% | Score 24.8 | DB 4 | Length 1900 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 53.0% | Pred. No. 27 | | |
| Matches 53 | Conservative 0 | Mismatches 47 | Indels 0 | Gaps 0 |
| QY | 1 | ctgtcccgagaggccccattcccaagctggagctgagtaagaaactcgagaaaccgac | 60 | |
| | | | | |
| Db | 551 | ctgtgtcttcacaggggtccacacacgtacttggcttcgcattctgcgaagtcccccacacacag | 492 | |
| QY | 61 | gcccaagatcaatcagcttcattgagcagagacacacctct | 100 | |
| | | | | |
| Db | 491 | gctgccttccttctccgcattatagagcccccacacacagct | 452 | |

Search completed: October 3, 2002, 16:24:51
Job time: 17131 sec


```

/protein_id="CAD10598.1"
/db_xref="GI:16608834"
/translation="MEPGLEHALRRTPSWSLGSEHQEENSFLEOENSSGNPSPAVTS
SSEVETGCRRAKRALRMTROKSAVEEGEPGEGRSPAAESTLXETFPPTTRALCA
DPADGTPPTGMDCLPSDCTASAGSGSTDDVLEATETPAWCELEGLLEEPALC
LSPPAPPPKLCMDLELKRRCQAQIYMRMEOETTCIDMAATSKLVIETDMELEKAFPA
LVAGVRAAPLWDSKOSFVGMLTITDFLVLHRRYSPLQVIEIDQHKETWRELY
LQGCFKPLVLSIPNDSLFEAVYTLIKNRHRLPVLDPVSGVNLHILHKRLKFLTF
GSLPRPSFLVLTQDGLGTFRLDVALEAPLITGLTDLFDVDRVSAALPVNNGGVG
VGLSRDFVHLAAQOTYNIHLDMSVGEALFQRTALCYSLSCOPHESLGEYDIRIANE
QVHRLVLYVDETQHLGLVSLSDILQALVLSDPAGIDALGA"

```

[illegible]

| RESULT | 2 | | | | |
|------------|---|-------------|-----|--------|-----------------|
| AX099776 | | | | | |
| LOCUS | AX099776 | 2109 bp | DNA | Linear | PAT 02-APR-2001 |
| DEFINITION | Sequence 3 from Patent WO0120003. | | | | |
| ACCESSION | AX099776 | | | | |
| VERSION | AX099776.1 | GI:13538810 | | | |
| KEYWORDS | . | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2109) | | | | |
| AUTHORS | Andersson, L., Looft, C., Kalm, E., Milan, D., Robic, A., Rogel-Galliard, C., Iannuccielli, N., Gellin, J., le Roy, P. and Chardon, P. | | | | |
| TITLE | Variants of the gamma chain of ampk, dna sequences encoding the same, and uses thereof | | | | |
| JOURNAL | Patent: WO 0120003-A 3 22-MAR-2001; INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ; Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE) | | | | |
| FEATURES | location/Qualifiers | | | | |

```

Query Match      100.0%; Score 101; DB 6; Length 2109;
Best Local Similarity 100.0%; Pred. No. 1e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 gctcaagttcctgacatcttgggtccctgcgcgcgcgcgcctctctctacgcac 60
|||||

```

| | | | |
|------------|------------------------------------|--|----------------------------|
| Db | 900 | GCTCAGATTCCGTGCACATCTTTGGTTCCTCCTGCGCCGCCCTTCCTCTACCGCAC | 959 |
| Oy | 61 | tatccaagatttgcgacatccgacatcctcggaacttgctg | 101 |
| | | | |
| Dd | 960 | TATCCAAGATTGGGCATCGGCACATTCGGAAGACTTGCTG | 1000 |
| RESULT | 3 | | |
| AX099802 | | | |
| LOCUS | AX099802 | 2115 bp | DNA linear PAT 02-APR-2001 |
| DEFINITION | Sequence 29 from Patent WO0120003. | | |
| ACCESSION | AX099802 | | |
| VERSION | AX099802.1 | GI:13538836 | |

| | |
|-----------|---|
| TITLE | Characterization of the gamma chain of ampk, dna sequences encoding the same, and uses thereof |
| JOURNAL | PATENT : WO 0120003-A-29 22-MAR-2001 ;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
location/Qualifiers |
| FEATURES | |
| ORGANISM | Homo sapiens |
| COLOID | Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumariota; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| REFERENCE | 1 (bases 1 to 2115) |
| AUTHORS | Andersson, L., Looft, C., Kalm, E., Milan, D., Robic, A.,
Rogel-Galliard, C., Iannuccielli, N., Gellin, J., le Roy, P. and
Chardon, P. |

| | |
|------------|--|
| CDS | /organism="Homo sapiens"
/db_xref="taxon:9606"
1..1395
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35801.1"
/db_xref="GI:13538837"
/translation="MSFLEQENSSMPSPAVYSSSERIRGKKRAKALRMTROKSYEEG
EPFGQEGPSPRTASTGLEGATFPPTTPIADQDPAGVSTPTGMCGLSDSTASAG
SSNDDELATLEPATTEAMECELEGLLEERPAFLCSQAPRPKLGMDLERKGAQTYM
RPMQDHCYDAMATSSKLVIFPFMLEIEKNAFPAVLVANGRAAPLIMSKSKOSFQMTLI
TDFLVLHRRYRSLPVOIYEIOKHLETMEETYLQCCFPPLVSIIPNDSLFPAVYTLI
KNKIHRLPYLVDPSGVNVLHTLTHKRLKLEHIFGSLPLPRPSFLYRTIDOLGIGTFPDL
AVYLERAPLITLADIFYDRAVSKALPVNKEGVGVGLTSRFDVTHLAQDTYHNLDSV
GEMLRQTTLEGVLSQPHESLGEVIDIRAREQVRLVLVDETQHLGLGVSSLDILO
ALVSPGIDILGALA" |
| BASE COUNT | 460 a 622 c 562 g 471 t |
| ORIGIN | |

| Query Match | 100.0% | Score 101 | DB 6 | Length 2115 |
|-----------------------|---|--|----------|-------------|
| Best Local Similarity | 100.0% | Pred. No. 1e-20 | | |
| Matches 101 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1 | gctcaagttcctgcacatccttggtccctgcgcgcgcgcgcctctctctacgcac | 60 | |
| Db | 906 | gctcaagttcctgcacatcttggttccctctcgcgcgcgcgcctctctctacgcac | 965 | |
| QY | 61 | tatccaaagatttggcattcggcacattccgagacttgcctt | 101 | |
| Db | 966 | tatccaaagatttggcattcggcacattccgagacttgcctt | 1006 | |
| RESULT | 4 | | | |
| AF214519 | | | | |
| LOCUS | AF214519 | 2115 bp | mRNA | linear |
| DEFINITION | Homo sapiens AMP-activated protein kinase gamma subunit (PRAKG3) | | | |
| ACCESSION | AF214519 | | | |
| VERSION | AF214519.1 | GT:8215661 | | |
| KEYWORDS | human. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |

REFERENCE 1 (bases 1 to 2115)
AUTHORS Milan.D., Jeon.J.T., Looft.C., Amarger.V., Robic.A., Thelander.M., Rogel-Galliard.C., Paul.S., Iannuccielli.N., Rask.L., Ronne.H., Lundstrom.K., Reinsch.N., Gellin.J., Kalm.E., Roy.P.L., Chardon.P. and Andersson.L.
JOURNAL MEDLINE 20280150
PUBMED 10818001
TITLE A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle
REFERENCE 2 (bases 1 to 2115)
AUTHORS Milan.D., Jeon.J.T., Looft.C., Amarger.V., Robic.A., Rogel-Galliard.C., Paul.S., Gellin.J., Lundstrom.K., Reinsch.N., Kalm.E., Le Roy.P., Chardon.P. and Andersson.L.
JOURNAL MEDLINE 20280150
PUBMED 10818001
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish University of Agricultural Sciences, BMC box 597, Uppsala 751 24, Sweden

FEATURES
source location/Qualifiers
1..2115 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p"
/tissue_type="skeletal muscle"
1..2115 /gene="PRKAG3"
1..1395 /gene="PRKAG3"
/note="AMPK3"
/codon_start=1
/product="AMP-activated protein kinase gamma subunit"
/protein_id="AA173987.1"
/db_xref="GI:8215682"
/translation="MSFLEQSSSPSPAYTSSEIRIGKRAKALRWTRKSYEEGEPGEGEPSPRSTAEETGLETATPXTKTLADADPAGCTPTGDCLPSTCTASAAGSSDDVELATERPATFEAMCELEKLEERRALCSQAPFPKILMDLKRPGAOITMRFMDHCTYDAAATSSKLVIFDTMLEIKKAFALVANGVRAAPLMDSKKOSFVGLTI TDFILVLRHYRSPVLOIYEI EOHK IETWREIYLOGCPPLVSI SPNDSLFEAVYTLI KNRIHRLPVLDPVSGNVLA I LTHKRLFLI FGSILRPSELYRT IODLIGIFERDL AVLEETAPILTALDLPVDRVSALPVNVECOVGLYRFPVITHLAAGQTYNHLDMVGEALRQRTLCLEGVLSQPHSEISGEVIDRIAREQVHRVLVDETHLGVYSLSDILOALVSPATIDALGA"

BASE COUNT 460 a 622 c 562 g 471 t
ORIGIN

Query Match 100.0% Score 101; DB 9; Length 2115;
Best Local Similarity 100.0% Pred. No. 1e-20;
Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctcaagtcctgcacacatcttgctcctgcctgcctgcctcctcctcctacgcac 60
|||||
DB 906 GCTCAAGTTCCTGCACATCTTGCTCCCTGCTGCCGCCCTCTCTACCGCAC 965
|||||

QY 61 tatccaagattggcgatcgcacatcgcagacttgctg 101
|||||
DB 966 TATCCAGATTGGGCGATCGCACATTCGAGACTTGGCTG 1006
|||||

RESULT 5
HSA249977
LOCUS HSA249977 2290 bp mRNA linear PRI 07-APR-2000
DEFINITION Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit (AMPK gamma 3 gene).
ACCESSION AJ249977
VERSION AJ249977.1 GI:6688200
KEYWORDS AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2290)
AUTHORS Cheung.P.C., Salt.I.P., Davies.S.P., Hardie.D.C. and Carling.D.
TITLE Characterization of AMP-activated protein kinase gamma-subunit isoforms and their role in AMP binding
JOURNAL Biochem. J. 346 Pt 3, 659-669 (2000)
MEDLINE 20164049
PUBMED 10818001
TITLE Direct Submission
JOURNAL Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC Clinical Sciences Centre, Hammersmith Hospital, DuCane Road, London, W12 0NN, UNITED KINGDOM

FEATURES
source location/Qualifiers
1..2290 /organism="Homo sapiens"
/db_xref="taxon:9606"
22..1500 /gene="AMPK gamma 3"
22..1500 /gene="AMPK gamma 3"
/function="AMP-activated protein kinase regulatory subunit"
/codon_start=1
/evidence="experimental"
/product="AMP-activated protein kinase gamma 3 subunit"
/protein_id="CAB65117.1"
/db_xref="GI:6688201"
/translation="MEDELHALRRTPSMSSLGGSEHOEMSFLEQSSSPSPAYTSSEIRIGKRAKALRWTRKSYEEGEPGEGSPRSTAEETGLETATPXTKTLADADPAGCTPTGDCLPSTCTASAAGSSDDVELATERPATFEAMCELEKLEERRALCSQAPFPKILMDLKRPGAOITMRFMDHCTYDAAATSSKLVIFDTMLEIKKAFALVANGVRAAPLMDSKKOSFVGLTI TDFILVLRHYRSPVLOIYEI EOHK IETWREIYLOGCPPLVSI SPNDSLFEAVYTLI KNRIHRLPVLDPVSGNVLA I LTHKRLFLI FGSILRPSELYRT IODLIGIFERDL AVLEETAPILTALDLPVDRVSALPVNVECOVGLYRFPVITHLAAGQTYNHLDMVGEALRQRTLCLEGVLSQPHSEISGEVIDRIAREQVHRVLVDETHLGVYSLSDILOALVSPATIDALGA"

BASE COUNT 501 a 674 c 617 g 498 t
ORIGIN

Query Match 100.0% Score 101; DB 9; Length 2290;
Best Local Similarity 100.0% Pred. No. 1e-20;
Matches 101: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctcaagtcctgcacacatcttgctcctgcctgcctgcctcctcctcctacgcac 60
|||||
DB 1002 GCTCAAGTTCCTGCACATCTTGCTCCCTGCTGCCGCCCTCTCTACCGCAC 1061
|||||

QY 61 tatccaagattggcgatcgcacatcgcagacttgctg 101
|||||
DB 1062 TATCCAGATTGGGCGATCGCACATTCGAGACTTGGCTG 1102
|||||

RESULT 6
AX099774
LOCUS AX099774 1867 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 1 from Patent W00120003.
ACCESSION AX099774
VERSION AX099774.1 GI:13538808
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
REFERENCE 1 (bases 1 to 1867)
AUTHORS Andersson.L., Looft.C., Kalm.E., Milan.D., Robic.A., Rogel-Galliard.C., Iannuccielli.N., Gellin.J., Le Roy.P. and Chardon.P.
TITLE Variants of the gamma chain of ampk, dna sequences encoding the same, and uses thereof
JOURNAL Patent: WO 0120003-A 1 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ; Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)

FEATURES
source
1..1867
/organism="Sus scrofa"
/db_xref="taxon:9823"
472..1389
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35798.1"
/db_xref="GI:13538809"
/translation="MHFQEHCTDAMATSSKLVIFDTMLEIKKAFALVANGVRAAP
LSDSKQSFVGMVLTIDFLLVHRYRSPVLOIYEIEHKIETWREIYLQCFPLVS
ISPDLSFEAYVALIKNRILHRLPYLDVPSGAVLILTHKRLKFLHIGTLPRPSFL
YRTIIDLIGTFRLDVAVLEFAPILTALDIFVDRVSALPVNENGVGVGLYSRFDV
HLAAGOTYNHLDMNVGFAIRORTLCTEGVLSCQPHETLGEVIDRIVRQVIRLVLYDE
TQHLGVSLSDILOALVLSFAGIDALGA"

BASE COUNT 380 a 583 c 529 g 375 t
ORIGIN

Query Match 92.7% Score 93.6: DB 6: Length 1867;
Best Local Similarity 96.0% Pred. No. 1.9e-18;
Matches 96: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

QY 2 ctcaagttccgtcacatcttggtccctgctgcgcgcgcctctctctaccgcaact 61
|||||
DB 901 CTCAGTTCCTGCACATCTTTGGCACCTGCGCCGCCCTCTCTTACCGCAC 960
|||||
QY 62 atccaagattgggcacatccgcagactggctg 101
|||||
DB 961 ATCCAAGATTGGGCATCGCACATTCCGACACTTGCGCG 1000
|||||

RESULT 7
AF214520 1873 bp mRNA linear MAM 03-JUN-2000
LOCUS
DEFINITION Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
ACCESSION AF214520
VERSION AF214520.1 GI:8215683
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS
1 (bases 1 to 1873)
Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
Rogel-Galliard,C., Paul,S., Jannuccelli,N., Rask,L., Ronne,H.,
Lundstrom,K., Reinsch,N., Gellin,J., Kaim,E., Roy,P.L., Chardon,P.
and Andersson,L.
TITLE
A mutation in PRKAG3 associated with excess glycogen content in pig
skeletal muscle
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 1873)
Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
Rogel-Galliard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
Kaim,E., Le Roy,P., Chardon,P. and Andersson,L.
TITLE
Direct Submission
JOURNAL
Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
Sweden
FEATURES
source
1..1873
/organism="Sus scrofa"
/db_xref="taxon:9823"
/chromosome="15"
/map="15q"
/tissue-type="skeletal muscle"
1..1873
/gene="PRKAG3"
1..1395
/gene="PRKAG3"

/note="AMPK3"
/codon_start=1
/product="AMP-activated protein kinase gamma subunit"
/protein_id="AAF73988.1"
/db_xref="GI:8215684"
/translation="MSFLEGGESRSPWSPRAVTTSSERSHGDCNKRMTROEDVEEG
GPGPREGPOSRPVASTGOEATFPKATPLAQAAPLAEVONPPTREDDILPSDCAASAS
DSNTDHLDLGTFESASASAGDELGLVEKPAFCSPPEVLLPRLCWDELQKPGAOVTH
HFMOEHCTYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLSDSKQSFVGMVLT
IDFLLVHRYRSPVLOIYEIEHKIETWREIYLQCFPLVSISPDLSFEAYVALIKNR
ILHRLPYLDVPSGAVLILTHKRLKFLHIGTLPRPSFLYRTIIDLIGTFRLDVA
VLEFAPILTALDIFVDRVSALPVNENGVGVGLYSRFDVHLAAGOTYNHLDMNV
GEALRORTLCTEGVLSCQPHETLGEVIDRIVRQVIRLVLYDETHLGVSLSDILO
ALVLSFAGIDALGA"

BASE COUNT 382 a 580 c 535 g 376 t
ORIGIN

Query Match 92.7% Score 93.6: DB 4: Length 1873;
Best Local Similarity 96.0% Pred. No. 1.9e-18;
Matches 96: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

QY 2 ctcaagttccgtcacatcttggtccctgctgcgcgcgcctctctctaccgcaact 61
|||||
DB 907 CTCAGTTCCTGCACATCTTTGGCACCTGCGCCGCCCTCTCTTACCGCAC 966
|||||
QY 62 atccaagattgggcacatccgcagactggctg 101
|||||
DB 967 ATCCAAGATTGGGCATCGCACATTCCGACACTTGCGCG 1006
|||||

RESULT 8
AX099800 1873 bp DNA linear PAT 02-APR-2001
LOCUS
DEFINITION Sequence 27 from Patent WO0120003.
ACCESSION AX099800
VERSION AX099800.1 GI:13538834
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS
1 (bases 1 to 1873)
Andersson,L., Looft,C., Kaim,E., Milan,D., Robic,A.,
Rogel-Galliard,C., Jannuccelli,N., Gellin,J., Le Roy,P. and
Chardon,P.
TITLE
Variants of the gamma chain of ampk, dna sequences encoding the
same, and uses thereof
JOURNAL
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
PATENT: WO 0120003-A 27 22-MAR-2001;
Andersson, Leif (SE) ; Looft, Christian (DE) ; Kaim, Ernst (DE)
FEATURES
source
1..1873
/organism="Sus scrofa"
/db_xref="taxon:9823"
1..1395
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35800.1"
/db_xref="GI:13538835"
/translation="MSFLEGGESRSPWSPRAVTTSSERSHGDCNKRMTROEDVEEG
GPGPREGPOSRPVASTGOEATFPKATPLAQAAPLAEVONPPTREDDILPSDCAASAS
DSNTDHLDLGTFESASASAGDELGLVEKPAFCSPPEVLLPRLCWDELQKPGAOVTH
HFMOEHCTYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLSDSKQSFVGMVLT
IDFLLVHRYRSPVLOIYEIEHKIETWREIYLQCFPLVSISPDLSFEAYVALIKNR
ILHRLPYLDVPSGAVLILTHKRLKFLHIGTLPRPSFLYRTIIDLIGTFRLDVA
VLEFAPILTALDIFVDRVSALPVNENGVGVGLYSRFDVHLAAGOTYNHLDMNV
GEALRORTLCTEGVLSCQPHETLGEVIDRIVRQVIRLVLYDETHLGVSLSDILO
ALVLSFAGIDALGA"

BASE COUNT 382 a 580 c 535 g 376 t
ORIGIN

Query Match 92.7% Score 93.6; DB 6; Length 1873;
Best Local Similarity 96.0% Pred. No. 1.9e-18;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ctcaagttccgcacatcttgcttgcctgctgcgcgcgcctctctctctaccgact 61
|||||
Db 907 CTCAGTTCCTCGCACATCTTTGGCACCCCTGCTGCCCGCCCTCTCTCTACCGCACC 966
|||||

QY 62 atccaagattggcgcacatccgcagacttgctg 101
|||||
Db 967 ATCCAGATTGGGCATCGCACATTCGAGACTTGGCCG 1006
|||||

RESULT 9
AX099804 2022 bp DNA linear PAT 02-APR-2001
LOCUS AX099804
DEFINITION Sequence 31 from Patent WO0120003.
ACCESSION AX099804
VERSION AX099804.1 GI:13538838
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Pig.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
Rogel-Galliard,C., Iannuccielli,N., Geillin,J., Le Roy,P. and
Charidon,P.
Variants of the gamma chain of ampf, dna sequences encoding the
same, and uses thereof
Patent: WO 0120003-A 31 22-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
Location/Qualifiers
1. .2022
/organism="Sus scrofa"
/db_xref="taxon:9823"

BASE COUNT 412 a 623 c 593 g 394 t
ORIGIN

Query Match 92.7% Score 93.6; DB 6; Length 2022;
Best Local Similarity 96.0% Pred. No. 1.9e-18;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ctcaagttccgcacatcttgcttgcctgctgcgcgcgcctctctctaccgact 61
|||||
Db 1057 CTCAGTTCCTCGCACATCTTTGGCACCCCTGCTGCCCGCCCTCTCTACCGCACC 1116
|||||

QY 62 atccaagattggcgcacatccgcagacttgctg 101
|||||
Db 1117 ATCCAGATTGGGCATCGCACATTCGAGACTTGGCCG 1156
|||||

RESULT 10
AX281580 1722 bp DNA linear PAT 03-NOV-2001
LOCUS AX281580
DEFINITION Sequence 3 from Patent WO0177305.
ACCESSION AX281580
VERSION AX281580.1 GI:16608831
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (sites)
Andersson,L., Luthman,H. and Marklund,S.
TITLE
Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL
Patient: WO 0177305-A 3 18-OCT-2001;
Aresis AB (SE)
Location/Qualifiers
1. .1722
/organism="Homo sapiens"

BASE COUNT 321 a 504 c 534 g 363 t
ORIGIN

Query Match 78.4% Score 79.2; DB 6; Length 1722;
Best Local Similarity 91.3% Pred. No. 4.9e-14;
Matches 84; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 10 cctgcacatcttgcttgcctgctgcgcgcgcctctctctaccgactccaga 69
|||||
Db 1510 CCATCCTACACAGAGGTTCCCTGCTGCCCGCCCTCTCTCTACCGACTATCCACA 1569
|||||

QY 70 ttgggcacatccgcacatccgcagacttgctg 101
|||||
Db 1570 TTGGGCATCGCACATTCGAGACTTGGCCG 1601
|||||

RESULT 11
AC027416/c 152129 bp DNA linear HTG 07-JUN-2000
LOCUS AC027416/c
DEFINITION Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
unorderd pieces.
AC027416
VERSION AC027416.2 GI:8317289
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 152129)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-504G11
2 (bases 1 to 152129)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hags,B., Heatford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Londoald,P., Marquis,N.,
McCarthy,M., McKean,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talmas,J.,
Testfaye,S., Theodore,J., Titrill,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2000 this sequence version replaced gi:7342115.
ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/xw/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W18R
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L7458
Center clone name: 504_G.11

Summary Statistics
Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135376 bases at least Q40
Consensus quality: 143264 bases at least Q30
Consensus quality: 146503 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 149029; sum-of-contigs
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1005: contig of 1005 bp in length
1006 1105: gap of 100 bp
1106 2402: contig of 1297 bp in length
2403 2502: gap of 100 bp
2503 3823: contig of 1321 bp in length
3824 3923: gap of 100 bp
3924 5020: contig of 1097 bp in length
5021 5120: gap of 100 bp
5121 6161: contig of 1041 bp in length
6162 6261: gap of 100 bp
6262 7547: contig of 1286 bp in length
7548 7647: gap of 100 bp
7648 9983: contig of 2336 bp in length
9984 10083: gap of 100 bp
10084 12556: contig of 2473 bp in length
12557 12656: gap of 100 bp
12657 15043: contig of 2387 bp in length
15044 15143: gap of 100 bp
15144 17123: contig of 1980 bp in length
17124 17223: gap of 100 bp
17224 19466: contig of 2243 bp in length
19467 19566: gap of 100 bp
19567 21928: contig of 2362 bp in length
21929 22028: gap of 100 bp
22029 24319: contig of 2291 bp in length
24320 24419: gap of 100 bp
24420 27059: contig of 2640 bp in length
27060 27159: gap of 100 bp
27160 30170: contig of 3011 bp in length
30171 30270: gap of 100 bp
30271 33968: contig of 3698 bp in length
33969 34068: gap of 100 bp
34069 38179: contig of 4111 bp in length
38180 38279: gap of 100 bp
38280 42366: contig of 4087 bp in length
42367 42466: gap of 100 bp
42467 46365: contig of 3899 bp in length
46366 46465: gap of 100 bp
46466 51285: contig of 4820 bp in length
51286 51385: gap of 100 bp
51386 55871: contig of 4486 bp in length
55872 55971: gap of 100 bp
55972 60595: contig of 4624 bp in length
60596 60695: gap of 100 bp
60696 66595: contig of 5900 bp in length
66596 66695: gap of 100 bp
66696 73218: contig of 6523 bp in length
73219 73318: gap of 100 bp
73319 77115: contig of 3797 bp in length
77116 77215: gap of 100 bp
77216 85022: contig of 7807 bp in length
85023 85122: gap of 100 bp
85123 93314: contig of 8192 bp in length
93315 93414: gap of 100 bp
93415 101193: contig of 7779 bp in length

FEATURES
source
1. 152129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-504G11"
/clone_11b="RP11 Human Male BAC"
1. 1005
/note="assembly-fragment"
1106. 2402
/note="assembly-fragment"
2503. 3823
/note="assembly-fragment"
3924. 5020
/note="assembly-fragment"
5121. 6161
/note="assembly-fragment"
6262. 7547
/note="assembly-fragment"
7648. 9983
/note="assembly-fragment"
10084. 12556
/note="assembly-fragment"
12657. 15043
/note="assembly-fragment"
15144. 17123
/note="assembly-fragment"
17224. 19466
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
19567. 21928
/note="assembly-fragment"
22029. 24319
/note="assembly-fragment"
24420. 27059
/note="assembly-fragment"
27160. 30170
/note="assembly-fragment"
30271. 33968
/note="assembly-fragment"
34069. 38179
/note="assembly-fragment"
38280. 42366
/note="assembly-fragment"
42467. 46365
/note="assembly-fragment"
46466. 51285
/note="assembly-fragment"
51386. 55871
/note="assembly-fragment"
55972. 60595
/note="assembly-fragment"
60696. 66595
/note="assembly-fragment"
66696. 73218
/note="assembly-fragment"
73319. 77115
/note="assembly-fragment"
77216. 85022
/note="assembly-fragment"
85123. 93314
/note="assembly-fragment"
93415. 101193
/note="assembly-fragment"
101294. 113090
101194 101293: gap of 100 bp
101294 113090: contig of 11797 bp in length
113091 113190: gap of 100 bp
113191 123496: contig of 10306 bp in length
123497 123596: gap of 100 bp
123597 137837: contig of 14241 bp in length
137838 137938: gap of 100 bp
137938 152129: contig of 14192 bp in length.
Location/Qualifiers

| | | | | | |
|----|-----------------------|--|--------------|--------------|---------------------------------|
| | Query Match | 78.4%: | Score 79.2: | DB 2: | Length 196554; |
| | Best Local Similarity | 91.3%: | Pred. | No. 5.2e-14; | |
| | Matches | 84: | Conservative | 0: | Mismatches 8; Indels 0; Gaps 0; |
| Oy | 10 | cclgacacatcttggtccgcgtcgcgccggccccttcctctcacgcagcatatacaaga | 69 | | |
| | | | | | |
| Dd | 59610 | CCATCTTAAACCAAGGGTTTCCTGCTGCACCAGGCCCTCCTTCTTAACCGCATATTCAAAGA | 59551 | | |
| | | | | | |
| Oy | 70 | tltggcaticgcagcatctcgagacttggctga | 101 | | |
| | | | | | |
| Dd | 59550 | TTTGGGCATCGCACACTTCCAGACTTGCGCTG | 59519 | | |

| RESULT | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS |
|--------|------------|------------------------------------|-------------|----------|-----------------|
| 13 | AC009974.c | 206654 bp | DNA | linear | PRI 09-JAN-2002 |
| | AC009974 | Homo sapiens BAC clone Rpl1-459119 | from 2, | complete | sequence. |
| | AC009974 | AC009974.9 | GI:16799058 | | |
| | HTG. | human. | | | |

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 206654)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)

TITLE: The sequence of Homo sapiens BAC clone RP11-459119
JOURNAL: Unpublished (2001)
REFERENCE: 3 (bases 1 to 206854)
AUTHORS: Waterson, R. H.
TITLE: Direct Submission
JOURNAL: Submitted (08-SEP-1999) Genome Sequencing Center,

JOURNAL. Submitted (08-SEP-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

| | |
|-----------|--|
| REFERENCE | 4 (bases 1 to 206854) |
| AUTHORS | Waterston, R. H. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (08-NOV-2001) Genome Sequencing Center, Washington |

REFERENCE
5 (bases 1 to 206854)
MO 63108, USA
University School of Medicine, 4444 Forest Park Parkway, St. Louis,

AUTHORS Waterston, R.H.
TITLE Direct Submission
SUBMITTED 03-JAN-2002
JOURNAL Genome Sequencing Center, Washington

REFERENCE
6 (bases 1 to 206854)
MO 63108, USA
UNIVERSITY SCHOOL OF MEDICINE, 4444 FOREST PARK PARKWAY, ST. LOUIS,
MO 63108, USA

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Nov 8, 2001 this sequence version replaced gi:13431203.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0459119

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPEC-11 human BAC library was made from the blood of one male neonate, as described by Oogoezawa, K., Moon, P.-Y., Zhao, B., Frengen, E., Tanno, M., Catanesse, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.regen.com>) or Pletzer de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is Rp11-1077K22; the clone sequenced to the right is Rp11-64105. Actual start of this clone is at base position 1 of Rp11-559119; actual end is at base position 206854 of Rp11-459119.

Data from AC079810 and AC073128 was used to finish this clone, AC009974. Polymorphisms have been identified between AC073128 and AC009974. A single plasmid region exists between 38812-18993. An unresolved tandem in the HERV SVA exists between 184390-185163. PCR suggests that approximately 1700 bps are missing.

FEATURES
SOU

```

source
1. .20b834
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone_RP11-459119"
/clone_11b="RP11-11"
1. .37
/note="match to EST BE314060 (NID:g9134719)"
1. .37
/note="match to EST BI059713 (NID:g14467240)"
1. .37
/note="match to EST BF183086 (NID:g11061273)"
1. .37
/note="match to EST AL567345 (NID:g12920610)"
1. .37
/note="match to EST AW880850 (NID:g8042860)"
1. .37
/note="match to EST BF304755 (NID:g11251653)"
1. .37
/note="similar to Homo sapiens EST BI114348 (NID:g14565249)"
1. .37
/note="match to EST BG477625 (NID:g13409904)"
1. .37
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
3. .37
/note="match to EST BE908408 (NID:g10402954)"
4. .37
/note="match to EST AL670836 (NID:g4850567) wa04g10.x1"
164. .662
/note="match to EST C05773 (NID:g1502549)"
220. .221
/note="match to EST BG470047 (NID:g13402322)"
281. .344
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
281. .344

```

| | | |
|--------------|--|-----------|
| misc_feature | /note="match to EST BE908408 (NID:g10402954)" | 281..344 |
| misc_feature | /note="match to EST BF304755 (NID:g11251653)" | 281..344 |
| misc_feature | /note="match to EST BF183086 (NID:g11061273)" | 281..344 |
| misc_feature | /note="similar to Homo sapiens EST B1114348 (NID:g14565249)" | 281..344 |
| misc_feature | /note="match to EST BC477625 (NID:g13409904)" | 281..344 |
| misc_feature | /note="match to EST B1059713 (NID:g14467240)" | 281..344 |
| misc_feature | /note="match to EST AL567345 (NID:g12920610)" | 281..344 |
| misc_feature | /note="match to EST BE314060 (NID:g9134719)" | 281..344 |
| misc_feature | /note="match to EST A1670836 (NID:g4850567)" | 281..344 |
| misc_feature | /note="match to EST A8808050 (NID:g8042860)" | 294..344 |
| misc_feature | /note="match to EST BC470047 (NID:g13402322)" | 594..763 |
| misc_feature | /note="match to EST A8808050 (NID:g8042860)" | 594..763 |
| misc_feature | /note="match to EST BE314060 (NID:g9134719)" | 594..764 |
| misc_feature | /note="match to EST B1059713 (NID:g14467240)" | 594..763 |
| misc_feature | /note="match to EST BC477625 (NID:g13409904)" | 594..742 |
| misc_feature | /note="match to EST BE047599 (NID:g8364652)" | 594..763 |
| misc_feature | /note="match to EST BE908408 (NID:g10402954)" | 594..763 |
| misc_feature | /note="match to EST AA043371 (NID:g1521226)" | 594..763 |
| misc_feature | /note="match to EST A1670836 (NID:g4850567)" | 594..763 |
| misc_feature | /note="match to EST BF183086 (NID:g11061273)" | 594..763 |
| misc_feature | /note="similar to Homo sapiens EST B1114348 (NID:g14565249)" | 599..763 |
| misc_feature | /note="match to EST AL567345 (NID:g12920610)" | 617..1084 |
| misc_feature | /note="match to EST AA481361 (NID:g2210913)" | 622..763 |
| misc_feature | /note="match to EST A1860958 (NID:g5514574)" | 622..763 |
| misc_feature | /note="similar to Mus musculus EST A1196847 (NID:g3749453)" | 684..763 |
| misc_feature | /note="match to EST BG92568 (NID:g14396638)" | 962..1084 |
| misc_feature | /note="match to EST A1656812 (NID:g4740791)" | 967..1084 |
| misc_feature | /note="match to EST BE908408 (NID:g10402954)" | 967..1085 |
| misc_feature | /note="match to EST BF304755 (NID:g11251653)" | 967..1091 |
| misc_feature | /note="match to EST AA043371 (NID:g1521226)" | 967..1071 |
| Query Match | 78.4%; Score 79.2; DB 9; Length 206854; | |

[illegible]

```
Search completed: October 3, 2002, 14:50:23
Job time: 12143 sec
```

| | | | | |
|-----------------------|--------|-----------|--------|--------------|
| Query Match | 75.28; | Score 76; | DB 4; | Length 5888; |
| Best local similarity | 85.08; | Pred NO 4 | 2e-13. | |

Best Local Similarity 85.0%; Pred. NO. 4.7e-13;
Matches 85; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

DY 2 ccaagllcctgcacatccttlyllccctgtgccccgccttcctctctaacgcact 61
| | | | | | | | | | | | | | | | | | | | | |
Db 3265 caactgtcccaaccaccaaggcacacctctgcccccggcccttcctttctaacgcgacac 3324

```
QY      62 atccaagattcggcatcgccacatlcgagacttgctg 101  
         |||||  
Db     3325 ATCCAAGATTGGGCATCGGCACATTCCGAGACTTGGCCG 3364
```

RESULT 15
AF336381/C

| | | | | | |
|------------|---|-----------|-----|--------|-----------------|
| LOCUS | AF336381 | 227724 bp | DNA | linear | HTG 02-APR-2003 |
| DEFINITION | Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN | | | | |

PROGRESS ***, 3 unordered pieces.

| | |
|-----------|------------------|
| RECESSION | AF330381 |
| VERSION | AF336381.1 |
| KEYWORDS | GI:13507298 |
| | HTC: HTCS PHASE1 |

| KEYWORDS | FILE |
|----------|-------------|
| SOURCE | house mouse |
| ORGANISM | house mouse |

ORGANISM

REFERENCE
1 (bases 1 to 227724)

| AUTHORS | TITLE |
|--|---------------------------------|
| Rump, A., Hayes, C., Brown, S.D.M. and Rosenthal, A. | Mono-chromatic light scattering |

mouse chromosome 1 genomic sequence

| COOKING | REFERENCE | AUTHORS |
|---------|---------------------|---------|
| 2 | (bases 1 to 227724) | Burns & |

| AUTHORS | TITLE | DIRECT |
|-----------|-------|--------|
| KUMIP, A. | | Direct |

JOURNAL Submitted (17-JA

Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
* NOTE: This is a 'working draft' sequence. It currently

is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

| | | |
|---|--------|---------------------------------|
| * | 17869: | contig. of 17869 bp in length |
| * | 17870 | gap of unknown length |
| * | 17970 | contig. of 14777 bp in length |
| * | 32747 | gap of unknown length |
| * | 32847: | contig. of 194878 bp in length. |

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .227724 |

```
/organism="Mus musculus"  
/db_xref="taxon:10090"
```

```
/chromosome="1"
```

| BASE COUNT | 57663 a | 55423 c | 56238 g | 58199 t | 201 others |
|------------------------|---------|---------|---------|---------|------------|
| clone="PAC510; PAC457" | | | | | |

ORIGIN

| Query Match | Score | DB 2; | Length |
|-------------|-----------|-------|----------------|
| 70.38; | Score 71; | DB 2; | Length 227724; |

best local similarity 93.7%; Pseq.NO. 1.1/e-11;
Matches 74; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 23 ggtctccctgtctgccccggccctctctctctctaacgcacatctccagaatttgcatcgcc 82

[illegible]

83 acattccgagacttgctg 101

Db 5471 ACATTCCGAGATTGGCTG 5453

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:46 ; Search time 719.93 Seconds

(without alignments)
240.868 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100

Perfect score: 101

Sequence: 1 gcccaagttcttcgacatct.....cacatccgagacttgcctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: N.Geneseq_032802.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-------------|
| 1 | 101 | 100.0 | 547 | 22 | ABAO8485 |
| 2 | 101 | 100.0 | 1647 | 22 | AAH43685 |
| 3 | 101 | 100.0 | 2109 | 22 | AAH43685 |
| 4 | 101 | 100.0 | 2115 | 22 | AAH43685 |
| 5 | 93.6 | 92.7 | 1867 | 22 | AAH43685 |
| 6 | 93.6 | 92.7 | 1873 | 22 | AAH43685 |
| 7 | 93.6 | 92.7 | 2022 | 22 | AAH43685 |
| 8 | 79.2 | 78.4 | 1732 | 22 | AAH43685 |
| 9 | 30.8 | 30.5 | 1435 | 20 | AAH43685 |

| | | | | | | |
|----|------|------|-------|----|-----------|----------------------|
| 10 | 30.8 | 30.5 | 1467 | 23 | AAH43685 | DNA encoding novel |
| 11 | 30.8 | 30.5 | 2223 | 22 | AAH43685 | Human CDNA sequenc |
| 12 | 29.6 | 29.3 | 566 | 22 | AAH43685 | Mammalian vestibul |
| 13 | 29.6 | 29.3 | 1326 | 21 | AAH43685 | Arabidopsis thaliana |
| 14 | 29.6 | 29.3 | 2289 | 23 | ABH08199 | Drosophila melanog |
| 15 | 29.6 | 29.3 | 4363 | 23 | ABH08198 | Drosophila melanog |
| 16 | 29.4 | 29.1 | 200 | 24 | AAH17161 | Mouse melanocyte h |
| 17 | 29.4 | 29.1 | 1259 | 14 | AAH051226 | Mouse MSH-R gene. |
| 18 | 29.4 | 29.1 | 1260 | 19 | AAH051226 | Mouse melanocyte s |
| 19 | 29.4 | 29.1 | 1260 | 20 | AAH051226 | Mouse melanocyte s |
| 20 | 29.4 | 29.1 | 1260 | 24 | AAH051226 | Mouse melanocyte h |
| 21 | 29.4 | 29.1 | 5007 | 23 | ABH07647 | Drosophila melanog |
| 22 | 29.4 | 29.1 | 12403 | 23 | ABH07646 | Drosophila melanog |
| 23 | 29.4 | 29.1 | 4413 | 23 | ABH02825 | Drosophila melanog |
| 24 | 29.4 | 29.1 | 6765 | 23 | ABH02824 | Drosophila melanog |
| 25 | 29.4 | 29.1 | 7461 | 23 | ABH04284 | Drosophila melanog |
| 26 | 28.8 | 28.5 | 398 | 22 | AAH4424 | Novel human polyoma |
| 27 | 28.8 | 28.5 | 9979 | 22 | AAH4424 | Human immune/thema |
| 28 | 28.8 | 28.5 | 13673 | 22 | AAH4424 | Human immune/thema |
| 29 | 28.8 | 28.5 | 13673 | 22 | AAH4424 | Human immune/thema |
| 30 | 28.4 | 28.1 | 1682 | 21 | AAH4424 | Human pancreatic c |
| 31 | 28.2 | 27.9 | 635 | 22 | AAH4424 | Human reproductive |
| 32 | 28.2 | 27.9 | 2378 | 23 | ABH06405 | Drosophila melanog |
| 33 | 28.2 | 27.9 | 2530 | 21 | AAH4424 | Arabidopsis thaliana |
| 34 | 28.2 | 27.9 | 2588 | 23 | ABH1171 | Drosophila melanog |
| 35 | 28.2 | 27.9 | 2857 | 22 | AAH15945 | Human CDNA sequenc |
| 36 | 28.2 | 27.9 | 3238 | 23 | AAH15945 | DNA encoding novel |
| 37 | 28.2 | 27.9 | 4378 | 23 | ABH06404 | Drosophila melanog |
| 38 | 28.2 | 27.9 | 4732 | 23 | ABH06404 | Drosophila melanog |
| 39 | 28.2 | 27.7 | 301 | 23 | AAH4424 | CDNA 939 encoding |
| 40 | 28.2 | 27.7 | 737 | 22 | AAH4424 | Human CDNA clone (|
| 41 | 28.2 | 27.7 | 3125 | 22 | AAH4424 | Human CDNA clone (|
| 42 | 28.2 | 27.7 | 3156 | 19 | AAH4424 | T-cell surface ant |
| 43 | 28.2 | 27.7 | 3156 | 20 | AAH4424 | Human CD97 protein |
| 44 | 28.2 | 27.7 | 3156 | 21 | AAH4424 | Human CD97 protein |
| 45 | 28.2 | 27.7 | 12043 | 24 | AAH4424 | Human intercellula |

ALIGNMENTS

| | | |
|----------|---|-------------------------|
| RESULT 1 | ABAO8485 | standard; CDNA: 547 BP. |
| ID | ABAO8485 | standard; CDNA: 547 BP. |
| AC | ABAO8485 | standard; CDNA: 547 BP. |
| XX | ABAO8485 | standard; CDNA: 547 BP. |
| DT | 11-JAN-2002 | (first entry) |
| XX | 11-JAN-2002 | (first entry) |
| DE | Human AMP-activated protein kinase subunit homologue CDNA, SEQ ID NO:261. | |
| XX | Human AMP-activated protein kinase subunit homologue CDNA, SEQ ID NO:261. | |
| KW | Human; cytokine; cell proliferation; cell differentiation; growth factor; | |
| KW | hematopoiesis regulation; tissue growth; immunomodulator; activin; | |
| KW | inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis; | |
| KW | proliferation; metastasis; cancer; tumor; hematopoietic disorder; | |
| KW | myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; | |
| KW | chronic inflammatory condition; proliferative retinopathy; | |
| KW | atherosclerosis; coronary heart disease; arterial ischemia; | |
| KW | bone disorder; osteoporosis; vascular growth disorder; | |
| KW | tissue regeneration; wound healing; infection; immune disorder; | |
| KW | cell culture; drug screening; gene therapy; antiinflammatory; | |
| KW | antisthmatic; antiarthritic; hemostatic; antiarteriosclerotic; | |
| KW | cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial; | |
| KW | antifungal; vulnery; antitumor; ss. | |
| XX | antifungal; vulnery; antitumor; ss. | |
| XX | antifungal; vulnery; antitumor; ss. | |
| OS | Homo sapiens. | |
| OS | Homo sapiens. | |
| PN | WO200157188-A2. | |
| XX | WO200157188-A2. | |
| PD | 09-AUG-2001. | |
| XX | 09-AUG-2001. | |
| PF | 05-FEB-2001; 2001WO-US03800. | |
| XX | 05-FEB-2001; 2001WO-US03800. | |

| PR | 03-FEB-2000: 2000US-0496914. |
|----|---|
| PR | 27-APR-2000: 2000US-0560875. |
| XX | |
| PA | (HYSE-) HYSEQ INC. |
| XX | |
| PI | Tang YT, Liu C, Drmanac RT; |
| XX | |
| PI | WPI: 2001-457740/49. |
| DR | P-PSDB: ABB11241. |
| XX | |
| PT | Human proteins and DNA encoding sequences useful for preventing, |
| PT | treating or ameliorating a medical condition in a mammalian subject |
| PT | e.g. arthritis and cancer - |
| PS | Claim 1: Page 429; 1963pp; English. |
| XX | |
| CC | Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and |
| CC | sequences ABA08225-ABA09574 represent nucleic acids encoding them. The |
| CC | invention also relates to vectors and recombinant host cells comprising a |
| CC | nucleotide of the invention, methods of producing the novel polypeptides, |
| CC | antibodies against the polypeptides, methods of detecting the nucleotides |
| CC | or polypeptides in a sample, and methods of identifying compounds which |
| CC | bind to polypeptides of the invention. Although novel, many of the |
| CC | polypeptides of the invention have homology to known proteins, thereby |
| CC | giving an insight into their probable biological activities, and hence |
| CC | potential therapeutic applications. The polypeptides of the invention may |
| CC | have various activities, including cytokine, cell proliferation or cell |
| CC | differentiation activities; stem cell growth factor activity; |
| CC | hematopoiesis regulatory activity; tissue growth factor activity; |
| CC | immunomodulatory activity; activin- or inhibin-related activities; |
| CC | chemotactic or chemokinetic activities; haemostatic, thrombotic or |
| CC | thrombolytic activities; receptor or ligand activities; or may be |
| CC | involved in oncogenesis, cancer cell proliferation or metastasis. |
| CC | Depending on their biological activities, polypeptides and nucleotides of |
| CC | the invention are useful for preventing, treating or ameliorating medical |
| CC | conditions, e.g., by protein or gene therapy. Such conditions include |
| CC | cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell |
| CC | disorders), chronic inflammatory conditions (e.g., asthma or arthritis), |
| CC | proliferative retinopathy, atherosclerosis, coronary heart disease, |
| CC | arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal |
| CC | vascular growth. Polypeptides involved with tissue regeneration and |
| CC | repair (or nucleic acids encoding them) may be used to promote wound |
| CC | healing (e.g., of burns, incisions and ulcers), while those with |
| CC | immunomodulatory activities may be used in the treatment of viral, |
| CC | bacterial and fungal infections in addition to immune disorders. |
| CC | Polypeptides with growth factor activity may be used in cell cultures to |
| CC | promote cell growth. For example, such polypeptides may be used to |
| CC | manipulate stem cells in culture to give rise to neuroepithelial cells |
| CC | that can be used to augment or replace cells damaged by illness, |
| CC | autoimmune disease or accidental damage. The polypeptides and nucleotides |
| CC | may also be used in the diagnosis of the above conditions, and in drug |
| CC | screening techniques. The present sequence represents a cDNA encoding a |
| CC | novel human polypeptide of the invention. |
| XX | |
| SO | Sequence 547 BP: 112 A; 172 C; 133 G; 130 T; 0 other; |
| | |
| | Query Match 100.0%; Score 101; DB 22; Length 547; |
| | Best Local Similarity 100.0%; Pred. No. 7.9e-23; |
| | Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0 |
| QY | 1 gctcaagttccctcacacatttgatttccctgtgtgccggccgccttctctacaccgac 60 |
| | |
| DB | 263 gctcaagttccctcacacatttgatttccctgtgtgccggccgccttctctacaccgac 322 |
| | |
| OY | 61 tatccaagatttggcgatcgcgcacatcgcgagacttgctg 101 |
| | |
| DB | 323 tatccaagatttggcgatcgcgcacatcgcgagacttgctg 363 |
| | |

RESULT 2

AAHA43685

ID AAHA43685 standard; cDNA: 1647 BP.

AAHA43685

ID AAHA43685 standard; cDNA: 1647 BP.

| Accession | Gene | Location/Qualifiers |
|--|---------------|---------------------------|
| AAH43685 | PRKAG3 | 20..1489 |
| 21-JAN-2002 | (first entry) | /*tag= a |
| PRKAG3 | CDNA | /product= "PRKAG3" |
| Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant; metabolic disease; diabetes; obesity; substitution; ss. | | 230 |
| Homo sapiens. | | /*tag= b |
| | | /label= "C230G" |
| | | /note= "Causes P71A" |
| | | 559 |
| | | /*tag= c |
| | | /label= "T559C" |
| | | /note= "Silent variation" |
| | | 1037 |
| | | /*tag= d |
| | | /label= "C1037T" |
| | | /note= "Causes R340W" |
| MO200177305-A2. | | |
| 18-OCT-2001. | | |
| 06-APR-2001; 2001MO-SE00765. | | |
| 07-APR-2000; 2000OUS-195665P. | | |
| (AREX-) AREXIS AB. | | |
| Andersson L, Luthman H, Marklund S; | | |
| WPI: 2001-657170/75. | | |
| P-PSDB: Q0B47679. | | |
| New variants of human AMP-activated protein kinase gamma3 subunit associated with a metabolic disease e.g. diabetes or obesity and method for determining a risk estimate of diseases in subject by detecting the variant - | | |
| Disclosure; Fig 5; 25pp; English. | | |
| This sequence represents the full length cDNA encoding the human AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant, is useful in determining a risk estimate of a metabolic disease, such as diabetes or obesity, in a subject. The variation may occur in exons 3, 4 or 10. In exon 3, a variation may be a substitution of a G for a C at nucleotide 320, resulting in the amino acid substitution P71A; in exon 4, a variation may be a substitution of a T for a C at nucleotide 550; and in exon 10, a variation may be a substitution of a T for a C at nucleotide 1037, resulting in the amino acid substitution R340W. There may also be nucleotide variation in intron 6. The numbering of these variations is based on the full length cDNA as given, rather than on position 1 of the open reading frame. | | |
| Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other; | | |

| | | | | |
|-----------------------|----------------|-----------------|----------|-------------|
| Query Match | 100.0% | Score 101 | DB 22 | Length 1647 |
| Best Local Similarity | 100.0% | Pred. No. 1e-22 | | |
| Matches 101 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |

QY 1 gctcaagttcctgcacatcttggttcctcgtgcgcgcgcgcctctctctaacgcac 60


```

Db 1000 gctcaagtcctgcacatcttggcttgcctgcgcgcgcgcctctctctaccgcac 1059
      |||
Qy 61 tatccaagattggcgcacatccgcagagactggctg 101
      |||
Db 1060 tatccaagattggcgcacatccgcagagactggctg 1100
      |||

RESULT 3
AAD03296
ID AAD03296 standard; DNA: 2109 BP.
AC AAD03296:
UT 13-JUN-2001 (first entry)
XX
DE Human AMPK gamma subunit muscle-specific isoform, PRKAG3 CDNA.
XX
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH 5'UTR 1..471
FH /*tag= a
FH CDS 472..1389
FH /*tag= b
FH /*product= "Human Prkag3 protein"
FH 3'UTR 1390..2109
FH /*tag= c

MO200120003-A2.
XX
PD 22-MAR-2001.
XX
IN 11-SEP-2000; 2000MO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401386.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannucci N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI: 2001-244810/25.
DR P-PSDB: AAE00221.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy.
XX
PS Claim 12; Fig 2; 71pp; English.
XX
XX
XX The present sequence is a CDNA encoding human adenosine monophosphate
XX (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
XX PRKAG3. Mutation in Prkag3 results in an altered regulation of
XX carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
XX useful as therapeutic for treating carbohydrate metabolism disorders such
XX as diabetes, obesity, and disorders associated with muscle metabolism
XX such as myopathy and cardiovascular diseases, to modulate AMPK
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX and its functionally altered mutants are useful for the diagnostic
XX evaluation, genetic testing and prognosis of a metabolic disorder,
XX preferably a carbohydrate metabolism disorder. Primers that can detect
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

```

```

CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;

Query Match 100.0%; Score 101; DB 22; Length 2109;
Best Local Similarity 100.0%; Pred. No. 1; le-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctcaagtcctgcacatcttggcttgcctgcgcgcgcctctctaccgcac 60
      |||
Db 900 gctcaagtcctgcacatcttggcttgcctgcgcgcgcctctctaccgcac 959
      |||
Qy 61 tatccaagattggcgcacatccgcagagactggctg 101
      |||
Db 960 tatccaagattggcgcacatccgcagagactggctg 1000
      |||

RESULT 4
AAD03320
ID AAD03320 standard; cDNA; 2115 BP.
XX
AC AAD03320:
UT 13-JUN-2001 (first entry)
XX
DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 CDNA.
XX
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH 5'UTR 1..1395
FH CDS /*tag= a
FH /*product= "Human complete Prkag3 protein"

MO200120003-A2.
XX
PD 22-MAR-2001.
XX
IN 11-SEP-2000; 2000MO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401386.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannucci N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI: 2001-244810/25.
DR P-PSDB: AAE00223.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy.
XX
PS Claim 12; Page 65-68; 71pp; English.

```

XX The present sequence is a cDNA encoding human adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC useful as therapeutic for treating carbohydrate metabolism disorders such
CC as diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.

SO Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;

Query Match 100.0%; Score 101; DB 22; Length 2115;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gctcaagttcctgcacatcttggctcctgtcgtccgcgcctcctcctaccgcac 60
Db 906 gctcaagttcctgcacatcttggctcctgtcgtccgcgcctcctcctaccgcac 965

OY 61 tatccaagattgggcacatccgcacatccgcagacttgctg 101
Db 966 tatccaagattgggcacatccgcacatccgcagacttgctg 1006

RESUI.T 5

AAD03295
ID AAD03295 standard; cDNA; 1867 BP.

AC AAD03295;

DT 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KM PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KM genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KM chromosome 15; ss.

XX Sus scrofa.

OS Key Location/Qualifiers
FH 1..471
FT /*tag= a
FT 472..1389
FT /*tag= b
FT /product= "Sus scrofa PRKAG3 protein"
FT 1390..1867
FT /*tag= c

PN MO200120003-A2.

XX 22-MAR-2001.

PF 11-SEP-2000; 2000MO-EP09896.

PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gallard C;
PI Iannucelli N, Gellin J, Le Roy P, Chardon P;
DR WPI; 2001-244810/25.
XX P-PSDB: AAE00220.

Claim 12; Fig 2; 71pp; English.

CC The present sequence is a cDNA encoding pig adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.

SO Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;

Query Match 92.7%; Score 93.6; DB 22; Length 1867;
Best Local Similarity 96.0%; Pred. No. 2.4e-20;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 ctcaagttcctgcacatcttggctcctgtcgtccgcgcctcctcctaccgcact 61
Db 901 ctcaagttcctgcacatcttggctcctgtcgtccgcgcctcctcctaccgcac 960

OY 62 atccaagattgggcacatccgcacatccgcagacttgctg 101
Db 961 atccaagattgggcacatccgcacatccgcagacttgctg 1000

RESULT 6

AAD03319
ID AAD03319 standard; cDNA; 1873 BP.

AC AAD03319;

DT 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KM PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KM genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KM chromosome 15; ss.

```

OS   Sus scrofa.
XX   FH   Key      Location/Qualifiers
XX   FT   CDS      1..1395
XX   FT   /tag="a
XX   FT   /product="Sus scrofa complete prkag3 protein"
XX   PN   WO200120003-A2.
XX   PD   22-MAR-2001.
XX   PP   11-SEP-2000; 2000WO-EP09896.
XX   PK   10-SEP-1999; 99EP-0402236.
XX   PR   18-MAY-2000; 2000EP-0401388.
XX   PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX   PA   (ANDE/) ANDERSSON L.
XX   PA   (LOOF/) LOOFT C.
XX   PA   (KALM/) KALM E.
XX   PI   Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX   PI   Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX   DR   WP1: 2001-244810/25.
XX   DR   P-PSDB: AAE00222.
XX   PT   New variants of the gamma subunit of vertebrate adenosine
XX   PT   monophosphate-activated kinase for diagnosis or treatment of disorders
XX   PT   associated with energy metabolism such as diabetes, obesity, and
XX   PT   myopathy -
XX   PS   Claim 12: Page 62-64; 71pp; English.
XX   CC   The present sequence is a cDNA encoding pig adenosine monophosphate
XX   CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
XX   CC   complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
XX   CC   15. Mutation in prkag3 results in an altered regulation of carbohydrate
XX   CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX   CC   therapeutic for treating carbohydrate metabolism disorders such as
XX   CC   diabetes, obesity, and disorders associated with muscle metabolism
XX   CC   such as myopathy and cardiovascular diseases, to modulate AMPK
XX   CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX   CC   and its functionally altered mutants are useful for the diagnostic
XX   CC   evaluation, genetic testing and prognosis of a metabolic disorder.
XX   CC   Preferably a carbohydrate metabolism disorder. Primers that can detect
XX   CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX   CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
XX   CC   from the expression of a functionally altered allele of PRKAG3.
XX   CC   Transgenic animal and host cell transformed with PRKAG3 or a
XX   CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX   CC   screening compounds able to modulate AMPK activity. Nucleic acid
XX   CC   encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or
XX   CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
XX   CC   of PRKAG3 and is useful in gene therapy.
XX   SQ   Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;

Query Match      92.7%; Score 93.6; DB 22; Length 1873;
Best Local Similarity 96.0%; Pred. No. 2, 4e-20;
Matches 96: Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY   2   ctaagttcctgcacatcttggtctccctgctgccccggcccccttccctctacgcagc 61
OY   DB   907   ctcaagttcctgcacatcttggtcaccctgcctgccccggcccccttccctctacgcagc 966
OY   62   atccaagattggcagtcggcacatccgagacttgctg 101
OY   DB   967   atccaagattggcagtcggcacatccgagacttgctg 1006

```

RESULT 7

```

AAD03321
ID   AAD03321 standard; DNA; 2022 BP.
XX   AC   AAD03321;
XX   DT   13-JUN-2001 (first entry)
XX   DE   Sus scrofa PRKAG3 splice variant DNA.
XX   KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
XX   KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
XX   KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
XX   KW   cystathione beta synthase; CBS; cardiact; gene therapy; ds.
XX   OS   Sus scrofa.
XX   FH   Key      Location/Qualifiers
XX   FT   CDS      1..1345
XX   FT   /tag="a
XX   FT   /product="Sus scrofa prkag3 splice variant"
XX   PN   WO200120003-A2.
XX   PD   22-MAR-2001.
XX   PP   11-SEP-2000; 2000WO-EP09896.
XX   PK   10-SEP-1999; 99EP-0402236.
XX   PR   18-MAY-2000; 2000EP-0401388.
XX   PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX   PA   (ANDE/) ANDERSSON L.
XX   PA   (LOOF/) LOOFT C.
XX   PA   (KALM/) KALM E.
XX   PI   Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX   PI   Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX   DR   WP1: 2001-244810/25.
XX   DR   P-PSDB: AAE00222.
XX   PT   New variants of the gamma subunit of vertebrate adenosine
XX   PT   monophosphate-activated kinase for diagnosis or treatment of disorders
XX   PT   associated with energy metabolism such as diabetes, obesity, and
XX   PT   myopathy -
XX   PS   Claim 12: Page 69; 71pp; English.
XX   CC   The present sequence is pig adenosine monophosphate (AMP)-activated
XX   CC   kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
XX   CC   variant DNA. Prkag3 gene is located in the RN locus of chromosome 15.
XX   CC   Mutation in prkag3 results in an altered regulation of carbohydrate
XX   CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX   CC   therapeutic for treating carbohydrate metabolism disorders such as
XX   CC   diabetes, obesity, and disorders associated with muscle metabolism
XX   CC   such as myopathy and cardiovascular diseases, to modulate AMPK
XX   CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX   CC   and its functionally altered mutants are useful for the diagnostic
XX   CC   evaluation, genetic testing and prognosis of a metabolic disorder,
XX   CC   preferably a carbohydrate metabolism disorder. Primers that can detect
XX   CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX   CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
XX   CC   from the expression of a functionally altered allele of PRKAG3.
XX   CC   Transgenic animal and host cell transformed with PRKAG3 or a
XX   CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX   CC   screening compounds able to modulate AMPK activity. Nucleic acid
XX   CC   encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or
XX   CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
XX   CC   of PRKAG3 and is useful in gene therapy.
XX   SQ   Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other;

```


| Field | Key | Location/Qualifiers |
|-------|-----|---|
| FT | CDS | 265..1503 |
| FT | | /*tag= a |
| XX | | WO9856052-A2. |
| XX | | |
| XX | | 23-DEC-1998. |
| XX | | |
| XX | | 19-JUN-1998; 98WO-US12813. |
| XX | | 19-JUN-1997; 97US-0878989. |
| XX | | |
| XX | | (INCY-) INCYTE PHARM INC. |
| PA | | |
| PI | | Bandman O, Corley NC, Goli SK, Guegler KJ, Hillman JL; |
| PI | | Lat P, Shah P; |
| XX | | |
| DR | | WPI; 1999-080952/07. |
| DR | | P-PSDB; AAW68438. |
| XX | | |
| PT | | New disease associated protein kinases - used to stimulate cell |
| PT | | proliferation and to treat the immune response and cancer |
| XX | | |
| PS | | Claim 5; Page 66-67; 93pp: English. |

This cDNA sequence codes for human disease associated protein kinase DAPK-7 (see AAM84438). DAPK-7 cDNA was first identified in the PENITENT01 cDNA library using a computer search for amino acid alignments, and a consensus sequence was derived from the extended and overlapping Incyte clones 3075712/HENRNT01, 842220/PROSTU05, 1364747/SCORNO02, 145972 and 145802/PENITUT01 and 1479332/CORNOT02. DAPK-7 shows 73% homology with the human foetal liver AMPK gamma subunit (GI 1335656), and is associated with cDNA libraries which are immortalised or cancerous and show inflammatory or immune responses. The invention provides disease associated protein kinases DAPK-1 to DAPK-7 (see AAM84432-38) and cDNA clones encoding them (see AAO06831-36 and AAO06882), as well as expression vectors, host cells, agonists, antagonists and antibodies. The invention further provides uses of such products in the diagnosis, prevention and treatment of diseases associated with cell proliferation, especially cancer or an immune response (claimed). Conditions that may be treated include adult respiratory distress syndrome, allergies, asthma, arteriosclerosis, bronchitis, emphysema, hypersosinophilia, myocardial or pericardial inflammation, rheumatoid arthritis, Addison's disease, AIDS, anaemia, atherosclerosis, various diseases of the digestive system, atopic dermatitis, dermatomyositis, diabetes mellitus, glomerulonephritis, gout, Grave's disease, lupus erythematosus, multiple sclerosis, myasthenia gravis, osteoarthritis, osteoporosis, pancreatitis, polycystic kidney disease, polymyositis, scleroderma, Sjogren's syndrome, autoimmune thyroiditis, complications of cancer, extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal and helminthic infections, and trauma (disclosed). The DAPK nucleic acids are also used in a method for detection of DAPK expression levels in a biological sample.

Sequence 1435 BP; 421 A; 298 C; 331 G; 385 T; 0 other;

| | | | | |
|--------------------------|--------|----------------|-----------|--------------|
| Query Match | 30.58; | Score 30.8; | DB 20; | Length 1435; |
| Best Local Similarity | 57.18; | Pred. No. 2.3; | | |
| Matches 56; Conservative | 0; | Mismatches 42; | Indels 0; | Gaps 0 |

[illegible]

RESULT 10

| | | |
|----|---|--------------------------|
| AA | 84265 | |
| ID | AA84265 | standard; cDNA; 1467 BP. |
| XX | | |
| AC | AA84265; | |
| XX | | |
| DT | 13-FEB-2002 | (first entry) |
| XX | | |
| DE | DNA encoding novel human diagnostic protein #20069. | |
| XX | | |
| KW | Human; chromosome mapping; gene mapping; gene therapy; forensic; | |
| KW | food supplement; medical imaging; diagnostic; genetic disorder; ss. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | W0200175067-A2. | |
| XX | | |
| PD | 11-OCT-2001. | |
| XX | | |
| PF | 30-MAR-2001; 2001WO-US08631. | |
| XX | | |
| PR | 31-MAR-2000; 2000US-0540217. | |
| PR | 23-AUG-2000; 2000US-0649167. | |

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

PS Claim 1; SEQ ID No 20069; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

Sequence 1467 BP; 416 A; 327 C; 343 G; 381 T; 0 other;

| | | | | |
|-----------------------|--------|----------------|--------|----------------|
| Query Match | 30.5%; | Score 30.8; | DB 23; | Length 1467; |
| Best Local Similarity | 57.1%; | Pred. No. 2.3; | | |
| Matches | 56; | Conservative | 0; | Mismatches 42; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

2 ctcgaatctccgcagacatcttggcttccgcgcgcgcgcgcctctctctctacgacat 61
 715 ctcaagttccctccagccttttatgtctgtatatagcgaagctcgtctctcatgaagcagaac 774
 62 atccaagaattgggcatacgcgcacatctccgaagacttggc 99
 775 ctgatatgagccttgataatagaacagctatcccaacacattgc 812

| | |
|----------|---|
| AAH14839 | 11 |
| ID | AAH14839 standard; cDNA; 2223 BP. |
| XX | |
| AC | AAH14839; |
| XX | |
| DT | 26-JUN-2001 (first entry) |
| XX | |
| DE | Human cDNA sequence SEQ ID NO:12660. |
| XX | |
| KW | Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss |
| OS | Homo sapiens. |
| PN | EP1074617-A2. |
| PD | |
| XX | |
| PP | 07-FEB-2001. |
| PX | |
| PF | 28-JUL-2000; 2000EP-0116126. |
| XX | |
| PR | 29-JUL-1999; 99JP-0248036. |
| PR | 27-AUG-1999; 99JP-0300253. |
| PR | 11-JAN-2000; 2000JP-0118776. |
| PR | 02-MAY-2000; 2000JP-0183767. |
| XX | 09-JUN-2000; 2000JP-0241899. |
| PA | (HELI-) HELIX RES INST. |
| PI | Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; |
| PI | Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T; |
| DR | WPI: 2001-318749/34. |
| PT | Primer sets for synthesizing polynucleotides, particularly the 5602 |
| PT | full-length CDNs defined in the specification, and for the detection |
| PT | and/or diagnosis of the abnormality of the proteins encoded by the |
| PT | full-length CDNs - |
| XX | |
| PS | Claim 8; SEQ ID 12660; 2537pp + CD ROW; English. |
| XX | |
| CC | The present invention describes primer sets for synthesising 5602 |
| CC | full-length CDNs defined in the specification. Where a primer set |
| CC | comprises: (a) an oligo-dT primer and an oligonucleotide complementary |
| CC | to the complementary strand of a polynucleotide which comprises one of |
| CC | the 5602 nucleotide sequences defined in the specification, where the |
| CC | oligonucleotide comprises at least 15 nucleotides; or (b) a combination |
| CC | of an oligonucleotide comprising a sequence complementary to the |
| CC | complementary strand of a polynucleotide which comprises a 5'-end |
| CC | sequence and an oligonucleotide comprising a sequence complementary to a |
| CC | polynucleotide which comprises a 3'-end sequence, where the |
| CC | oligonucleotide comprises at least 15 nucleotides and the combination of |
| CC | the 5'-end sequence/3'-end sequence is selected from those defined in |
| CC | the specification. The primer sets can be used in antisense therapy and |
| CC | in gene therapy. The primers are useful for synthesising polynucleotides, |
| CC | particularly full-length CDNs. The primers are also useful for the |
| CC | detection and/or diagnosis of the abnormality of the proteins encoded by |
| CC | cDNAs easily without any specialised methods. AAH03166 to AAH13678 and |
| CC | AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to |
| CC | AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 |
| CC | represent oligonucleotides, all of which are used in the exemplification |
| CC | of the present invention. |
| XX | |
| Sequence | 2223 BP; 633 A; 431 C; 496 G; 663 T; 0 other; |

| | | | | |
|-----------------------|--------|--|--------|----------------|
| Query Match | 30.5% | Score 30.8; | DB 22; | Length 2223; |
| Best Local Similarity | 57.1%; | Pred. No. 2.5; | | |
| Matches | 56; | Conservative | 0; | Mismatches 42; |
| | | | Indels | 0; |
| | | | Gaps | 0; |
| 0y | 2 | ctcaagttctgcgaacattcttggtctctctgcgcgcgcgcctctctctctcagcact | 61 | |

Dy 62 atccaagatttgcgcatcgcacatlcgcgaacttcg 99
| | | | | | | | | | | | | |
Db 709 ctgtgatgacttgtaataatagaacatccacaacattgc 746

```

RESULT 12
AAS05559/c
ID AAS05559 standard; DNA; 566 BP.

```

DT 07-SEP-2001 (first entry)

Mammalian vestibular system geotactic behaviour modulator gene #159.

| | |
|----|---|
| KM | maximalist vestibular system; invertebrate; geotactic behaviour; vertigo. |
| KM | graviperceptive disorder; motion sickness; labyrinthitis; syphilis; ds. |
| KM | Meniere's disease; acoustic neuroma; multiple sclerosis; epilepsy; |
| KM | trauma; infection of the middle ear; ototoxic agent exposure. |

PN W0200140519-A2.

PD 07-JUN-2001.

01-DEC-2000; 2000WO-US32639.

PR 02-DEC-1999; 99US-0168579.

26-SEP-2000; 2000US-0669751.

(NEUR-) NEUROSCIENCES RES FOUND INC.

PI Greenspan RJ;

DR WPI; 2001-356159/37.

PT New isolated nucleic acid having mammalian vestibular system-modulating activity useful in the treatment of disorders such as motion sickness and vertigo -

PS Claim 59; Page 144; 179pp; English.

The sequences shown in AAS05401-AAS05661 represent DNA with mammalian vestibular system-modulating activity. The DNA sequences can be used in a method whereby a first and second strain of an invertebrate is obtained, and both are subjected to conditions in which the strains exhibit different geotactic behaviour. Genes that are differentially expressed in the first strain relative to the second strain are then identified. Mammalian genes having substantially the same nucleic acid sequence as these genes are used to decrease the symptoms of gravityperceptive disorders, such as motion sickness, vertigo, labyrinthitis, Meniere's disease, acoustic neuroma, multiple sclerosis, sphyllitis, trauuma, infection of the middle ear, exposure to ototoxic agents and epilepsy.

Sequence 566 BP; 191 A; 129 C; 161 G; 85 T; 0 other;

| | | | | |
|--------------------------|--------|----------------|-----------|-------------|
| Query Match | 29.38; | Score 29.6; | DB 22; | Length 566; |
| Best Local Similarity | 68.38; | Pred. No. 4.4; | | |
| Matches 41; Conservative | 0; | Mismatches 19; | Indels 0; | Gaps 0; |

27 13 gcaacatcttggttccctgctgcccccgcctccttcctctaccgcatataccaagatt 72
 || ||| | | | | | | | | | | | | | | |
544 GCTCATCAAGGTGACTACCTGTTCTCGCTTCTTCTCCTACTCATTTATCCCTGATCT 485

RESULT 13
AAC37238

ID AAC37238 standard: DNA: 1326 BP.
XX AAC37238;
AC
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16667.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PM EPI033405-A2.
PD
XX
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135253.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```

PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150864.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      29.3%; Score 29.6; DB 21; Length 1326;
Best Local Similarity 64.7%; Pred. No. 5.4;
Matches 44; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

```

QY 8 ttctcgacacatttggtctccgtgcccgcctctctctctacgcacatacaa 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 193 ttctcgacacatttggtctccgtgcccgcctctctctacgcacatacaa 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 gatttgagg 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 253 gtttaggg 260

```

```

RESULT 14
ABL08199/c
ID ABL08199 standard; cDNA: 2289 BP.
XX
AC ABL08199;

```

```

XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 19079.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX WPI. 2001-656860/75.
XX P-PSDB; ABB64096.
XX
DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1: SEQ ID NO 19079; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABE57737-ABE72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2289 BP; 694 A; 541 C; 589 G; 465 T; 0 other;

```

```

Query Match      29.3%; Score 29.6; DB 23; Length 2289;
Best Local Similarity 68.3%; Pred. No. 6.1;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 13 gacatcttggtccctgctgcccgcctctctctacgcacatacaaatlt 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 544 GCTCATCAAGGTGACTACCTGTTCTCTCTCTCTACCTATTCCTGATCT 485

```

```

RESULT 15
ABL08198/c
ID ABL08198 standard; cDNA: 4363 BP.
XX
AC ABL08198;
XX

```

26-MAR-2002 (first entry)

Drosophila melanogaster expressed polynucleotide SEQ ID NO 19076.

Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ss.

Drosophila melanogaster.

WO200171042-A2.

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venler JC, Adams M, Li PWD, Myers EW;
XX
UR WPI; 2001-656860/75.
DR P-PSDB; ABB64095.
XX
PI New isolated nucleic acid detection reagent for detecting 1000 or more
PI genes from Drosophila and for elucidating cell signalling and cell-cell
PI interactions -
XX
PS
PS Claim 1: SEQ ID NO 19076; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences and pharmaceutical drugs. The invention
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 4363 BP; 1461 A; 939 C; 974 G; 989 T; 0 other;

Query Match 29.3%; Score 29.6; DB 23; Length 4363;
Best Local Similarity 68.3%; Pred. No. 7;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 13 gcaatcttggtccctgctgcgcgcgcctcctcctctactacgaactatccaagattt 72
II IIII III I III I III I III III III I III III I
Db 1544 GCTCATCAAGTCACTACCTGTTCTTCCTTCCTCTCTACCTCATATCCCTGATCT 1485

Search completed: October 3, 2002, 16:30:48
Job time: 14318 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:08 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Perfect score: 101
Sequence: 1 gcccaattctctgcacatc.....cacattccagacttgctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 93.6 | 92.7 | 572 | 10 | B1344527 |
| 2 | 59 | 58.4 | 413 | 9 | AA178898 |
| 3 | 48.8 | 48.3 | 633 | 10 | B1072114 |
| 4 | 38.8 | 38.4 | 514 | 10 | BC037921 |
| 5 | 34.8 | 34.5 | 918 | 10 | BC027175 |
| 6 | 32.2 | 31.9 | 451 | 9 | AM147353 |
| 7 | 32.2 | 31.9 | 470 | 10 | B1036397 |
| 8 | 32.2 | 31.9 | 580 | 9 | AV914030 |
| 9 | 31.4 | 31.1 | 684 | 10 | BC101440 |
| 10 | 31.2 | 30.9 | 536 | 10 | BM488662 |
| 11 | 31.2 | 30.9 | 595 | 10 | BM487789 |
| 12 | 31.2 | 30.9 | 647 | 10 | BM440762 |
| 13 | 31.2 | 30.9 | 649 | 9 | AJ395115 |
| 14 | 31.2 | 30.9 | 758 | 9 | AJ396118 |
| 15 | 30.8 | 30.5 | 413 | 10 | H64260 |
| 16 | 30.8 | 30.5 | 451 | 10 | W39604 |
| 17 | 30.8 | 30.5 | 474 | 10 | BE803695 |

| | | | | | |
|----|------|------|-----|----|----------|
| 18 | 30.8 | 30.5 | 506 | 10 | B1760420 |
| 19 | 30.8 | 30.5 | 525 | 10 | BC148719 |
| 20 | 30.8 | 30.5 | 601 | 10 | BE846246 |
| 21 | 30.8 | 30.5 | 764 | 9 | AU138307 |
| 22 | 30.8 | 30.5 | 849 | 10 | B1657718 |
| 23 | 30.8 | 30.5 | 886 | 10 | BC572824 |
| 24 | 30.8 | 30.5 | 909 | 9 | AL558594 |
| 25 | 30.8 | 30.5 | 951 | 9 | AL513784 |
| 26 | 30.8 | 30.5 | 967 | 9 | AL524822 |
| 27 | 30.4 | 30.1 | 312 | 10 | C06172 |
| 28 | 30.2 | 29.9 | 924 | 12 | CNS029MX |
| 29 | 30 | 29.7 | 663 | 10 | BC415737 |
| 30 | 29.8 | 29.5 | 784 | 9 | AU080009 |
| 31 | 29.8 | 29.5 | 964 | 10 | BC035134 |
| 32 | 29.6 | 29.3 | 322 | 9 | AV525978 |
| 33 | 29.6 | 29.3 | 619 | 10 | BE978092 |
| 34 | 29.6 | 29.3 | 626 | 10 | BF528264 |
| 35 | 29.6 | 29.3 | 715 | 10 | BF488250 |
| 36 | 29.6 | 29.3 | 754 | 9 | A1063338 |
| 37 | 29.6 | 29.3 | 836 | 10 | B1659310 |
| 38 | 29.4 | 29.1 | 388 | 12 | AO535079 |
| 39 | 29.4 | 29.1 | 559 | 9 | A1062936 |
| 40 | 29.4 | 29.1 | 559 | 9 | A1062937 |
| 41 | 29.4 | 29.1 | 588 | 10 | BM340317 |
| 42 | 29.4 | 29.1 | 699 | 12 | BM447226 |
| 43 | 29.4 | 29.1 | 888 | 9 | AL581895 |
| 44 | 29.2 | 28.9 | 224 | 9 | AA045087 |
| 45 | 29.2 | 28.9 | 388 | 10 | B1774231 |

ALIGNMENTS

RESULT 1
B1344527 572 bp mRNA linear EST 30-JUL-2001
DEFINITION 373008 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
LOCUS B1344527
VERSION B1344527.1 GI:15037807
ACCESSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE

1 (bases 1 to 572)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keeler,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

TITLE

JOURNAL

Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.

PCR PRIMERS

FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCACAGTCACACG
Plate: 119 row: 1 column: 11
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta.
BASE COUNT 106 a 190 c 151 g 124 t 1 others
ORIGIN

Query Match 92.7%; Score 93.6; DB 10; Length 572;
Best Local Similarity 96.0%; Pred. No. 7.9e-16;
Matches 96; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 2 ctcaagcttcgcacatcttggtccctgcgcgcgcctctctctacgcact 61
|||||
Db 163 CTCACGCTCCGACATCTTGGCACCCCTGCGCCGCCCTCTCTACCGCACC 222
CTCACGCTCCGACATCTTGGCACCCCTGCGCCGCCCTCTCTACCGCACC 222
OY 62 atccaagattggcgcgcgcacatccgcagactgctg 101
|||||
Db 223 ATCCAGACTTTGGCGCATCGCACATTCGAGACTTGGCGCG 262
ATCCAGACTTTGGCGCATCGCACATTCGAGACTTGGCGCG 262

RESULT 2
AA178898 413 bp mRNA linear EST 09-MAR-1998
LOCUS 2938010.r1 StrataGene muscle 937209 Homo sapiens cDNA clone
DEFINITION IMAGE:611731 5' similar to SW:AAKQ_RAT P80385 5'-AMP-ACTIVATED
PROTEIN KINASE, GAMMA CHAIN ;, mRNA sequence.
ACCESSION AA178898
VERSION AA178898.1 GI:1760259
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Scheinberg, K., Stepien, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1280 Std Error: 0.00
Seq primer: -28M13 rev2 from AmerSham
High quality sequence stop: 255.
Location/Qualifiers
1. 413

FEATURES
source
/organism="Homo sapiens"
/db_xref="CDB:4643570"
/db_xref="taxon:9606"
/clone="IMAGE:611731"
/clone_lib="StrataGene muscle 937209"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skeletal muscle; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
Primer: Oligo dT. Skeletal muscle from patient with
malignant hyperthermia. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGCAGAC
3' -3' adaptor sequence: 5' CTCGACTTTTCTTTTCTTTTCTTTT 3' -
BASE COUNT 80 a 109 c 117 g 102 t 5 others
ORIGIN

Query Match 58.4%; Score 59; DB 9; Length 413;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 ctctctctctaccgcacatctccaagatttggcgcgcgcacatccgcagactgctg 101
|||||
Db 1 CTCCTTCCTCTACCGCAGCTATCCAGATTGGCGCATCGCACATTCGAGACTTGGCGCTG 59
CTCCTTCCTCTACCGCAGCTATCCAGATTGGCGCATCGCACATTCGAGACTTGGCGCTG 59

RESULT 3
BJ072114 633 bp mRNA linear EST 11-DEC-2001
LOCUS BJ072114 NIBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone X1096j16 5', mRNA sequence.
ACCESSION BJ072114
VERSION BJ072114.1 GI:17502303
KEYWORDS EST.
SOURCE African clawed frog.
Xenopus laevis
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 633)
Kitayama, A., Terashaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara,
Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-I
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. 633

FEATURES
source
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="X1096j16"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 144 a 162 c 148 g 179 t
ORIGIN

Query Match 48.3%; Score 48.8; DB 10; Length 633;
Best Local Similarity 68.0%; Pred. No. 0.0015;
Matches 68; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 2 ctcaagcttcgcacatcttggtccctgcgcgcgcctctctctacgcact 61
|||||
Db 131 CTCACGCTCCGACATCTTGGCACCCCTGCGCCGCCCTCTCTACCGCACC 190
CTCACGCTCCGACATCTTGGCACCCCTGCGCCGCCCTCTCTACCGCACC 190
OY 62 atccaagattggcgcgcgcacatccgcagactgctg 101
|||||
Db 191 ATCTGAGCTGGCGCATCGCACATTCAGAGATATAGCAG 230
ATCTGAGCTGGCGCATCGCACATTCAGAGATATAGCAG 230

RESULT 4
BG037921 514 bp mRNA linear EST 24-JAN-2001
LOCUS BG037921 NICHG XGC Emb1 Xenopus laevis cDNA clone IMAGE:3402446
DEFINITION 5' similar to TR:Q9ULX8 Q9ULX8 H91620P.;, mRNA sequence.
ACCESSION BG037921
VERSION BG037921.1 GI:12480506
KEYWORDS EST.
SOURCE African clawed frog.
Xenopus laevis
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 514)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

| | | | | |
|-----------------------|-----------------|----------------|-----------|-------------|
| Query Match | 31.9% | Score 32.2; | DB 10; | Length 470; |
| Best Local Similarity | 61.2% | Pred. No. 47; | | |
| Matches 52; | Conservative 0; | Mismatches 33; | Indels 0; | Gaps 0. |

| RESULT | 9 |
|------------|---|
| LOCUS | BG101440 |
| DEFINITION | BG101440 684 bp mRNA linear EST 29-JAN-2001
u72cc12.y1 McCarrey Rddy round spermatid Mus musculus cDNA clone
IMAGE:3664919.5 similar to TR:090LX8 Q90LX8 H91620P. ; mRNA
SEQUENCE. |
| ACCESSION | BG101440 |
| VERSION | BG101440.1 |
| KEYWORDS | GI:12596757 |
| SOURCE | EST. |
| ORGANISM | house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scleromorphi; Muridae; Murinae; Mus |

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:24:51 ; Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Sequence: 1 gctcaagctctctgacatct.....cacatccgagactgctg 101

Scoring table:

IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 30.8 | 30.5 | 1435 | 2 US-08-878-989-14 | Sequence 14, Appl |
| 2 | 30.8 | 30.5 | 1435 | 4 US-09-272-796-14 | Sequence 14, Appl |
| 3 | 29.4 | 29.1 | 1260 | 1 US-07-866-979-3 | Sequence 3, Appl |
| 4 | 29.4 | 29.1 | 1260 | 2 US-08-466-9068-3 | Sequence 3, Appl |
| 5 | 29.4 | 29.1 | 1260 | 3 US-08-706-281A-3 | Sequence 3, Appl |
| 6 | 29.4 | 29.1 | 1260 | 4 US-09-201-746-3 | Sequence 3, Appl |
| 7 | 29.4 | 29.1 | 1260 | 4 US-09-097-231-3 | Sequence 3, Appl |
| 8 | 27.8 | 27.5 | 1378 | 1 US-08-759-848-2 | Sequence 2, Appl |
| 9 | 27.8 | 27.5 | 1378 | 5 PCT-US95-09383-2 | Sequence 2, Appl |
| 10 | 27.6 | 27.3 | 5020 | 3 US-08-938-291A-3 | Sequence 3, Appl |
| 11 | 26.4 | 26.1 | 5975 | 1 US-08-336-257A-3 | Sequence 3, Appl |
| 12 | 25.6 | 25.3 | 1984 | 2 US-08-822-028-16 | Sequence 16, Appl |
| 13 | 25.6 | 25.3 | 1984 | 4 US-08-479-285-16 | Sequence 16, Appl |
| 14 | 25.4 | 25.1 | 1509 | 2 US-08-481-337A-1 | Sequence 1, Appl |
| 15 | 25.4 | 25.1 | 1509 | 2 US-08-696-2688-1 | Sequence 1, Appl |
| 16 | 25.4 | 25.1 | 1509 | 5 PCT-US95-05467-1 | Sequence 1, Appl |
| 17 | 25.4 | 25.1 | 1596 | 5 PCT-US94-11328A-3 | Sequence 1, Appl |
| 18 | 25.4 | 25.1 | 1702 | 4 US-08-934-254-26 | Sequence 26, Appl |
| 19 | 25.4 | 25.1 | 1984 | 4 US-09-382-256-1 | Sequence 1, Appl |
| 20 | 25.4 | 25.1 | 1984 | 4 US-09-395-115-1 | Sequence 1, Appl |
| 21 | 25.4 | 25.1 | 1984 | 4 US-08-436-265-1 | Sequence 1, Appl |
| 22 | 25.2 | 25.0 | 2264 | 4 US-09-126-109-9 | Sequence 1, Appl |
| 23 | 25.2 | 25.0 | 5962 | 6 5386025-5 | Sequence 1, Appl |
| 24 | 25.2 | 25.0 | 5975 | 1 US-08-404-354B-1 | Sequence 1, Appl |
| 25 | 25.2 | 25.0 | 5975 | 1 US-08-314-083B-1 | Sequence 1, Appl |
| 26 | 25.2 | 25.0 | 5975 | 1 US-08-435-675B-1 | Sequence 1, Appl |
| 27 | 25.2 | 25.0 | 5975 | 3 US-08-884-599-1 | Sequence 1, Appl |

| | | | | | |
|------|------|------|------|---------------------|------------------|
| C 28 | 24.8 | 24.6 | 1419 | 1 US-08-103-739B-1 | Sequence 1, Appl |
| C 29 | 24.8 | 24.6 | 1419 | 1 US-08-474-404-1 | Sequence 1, Appl |
| C 30 | 24.8 | 24.6 | 1419 | 2 US-08-485-845-1 | Sequence 1, Appl |
| C 31 | 24.8 | 24.6 | 1419 | 2 US-08-482-714-1 | Sequence 1, Appl |
| C 32 | 24.8 | 24.6 | 1419 | 4 US-09-211-416-1 | Sequence 1, Appl |
| C 33 | 24.8 | 24.6 | 1419 | 4 US-09-059-958-1 | Sequence 1, Appl |
| C 34 | 24.4 | 24.2 | 261 | 2 US-08-592-383-7 | Sequence 7, Appl |
| C 35 | 24.4 | 24.2 | 1071 | 1 US-08-612-986-1 | Sequence 1, Appl |
| C 36 | 24.4 | 24.2 | 1071 | 1 US-08-361-806A-1 | Sequence 1, Appl |
| C 37 | 24.4 | 24.2 | 1071 | 5 PCT-US95-16806A-1 | Sequence 1, Appl |
| C 38 | 24.4 | 24.2 | 1094 | 2 US-08-902-294-1 | Sequence 1, Appl |
| C 39 | 24.4 | 24.2 | 1094 | 3 US-09-178-637-1 | Sequence 1, Appl |
| C 40 | 24.4 | 24.2 | 2658 | 2 US-08-592-383-3 | Sequence 1, Appl |
| C 41 | 24.4 | 24.2 | 2928 | 2 US-08-095-728B-3 | Sequence 3, Appl |
| C 42 | 24.4 | 24.2 | 2928 | 5 PCT-US92-02320A-3 | Sequence 3, Appl |
| C 43 | 24.4 | 24.2 | 2940 | 2 US-08-592-383-1 | Sequence 3, Appl |
| C 44 | 24.4 | 24.2 | 2940 | 6 5171671-1 | Sequence 1, Appl |
| C 45 | 24.4 | 24.2 | 4002 | 1 US-08-331-488A-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-08-878-989-14
; Sequence 14, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Gali, Preeti
; APPLICANT: Gali, Surya K.
; APPLICANT: Shah, Purni
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PENITUT01

CLONE: 1452972
US-08-878-989-14

Query Match 30.5%: Score 30.8; DB 2; Length 1435;
Best Local Similarity 57.1%: Pred. No. 0.11;
Matches 56: Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 2 ctcaagttcctgacacatcttggtccctgcgcgcgcctctctctacgcact 61
DB 617 CTCAGTTCCTCCACACTTTTATGTCGATATGCCAAGCCTGCTTCATGAGCAGAAC 676
OY 62 atccaagattgggcatcgcgcacatccgcagactgac 99
DB 677 CTGATGAGCTTGGAATAGAACGTACCAACAACATTGC 714

RESULT 2
US-09-272-796-14
Sequence 14, Application US/09272796
Patent No. 6207148
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goll, Surya K.
APPLICANT: Shah, Puri
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/272,796
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/878,989
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PENITUT01
CLONE: 1452972
US-09-272-796-14

Query Match 30.5%: Score 30.8; DB 4; Length 1435;
Best Local Similarity 57.1%: Pred. No. 0.11;

Matches 56: Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 2 ctcaagttcctgacacatcttggtccctgcgcgcgcctctctctacgcact 61
DB 617 CTCAGTTCCTCCACACTTTTATGTCGATATGCCAAGCCTGCTTCATGAGCAGAAC 676
OY 62 atccaagattgggcatcgcgcacatccgcagactgac 99
DB 677 CTGATGAGCTTGGAATAGAACGTACCAACAACATTGC 714

RESULT 3
US-07-866-979-3
Sequence 3, Application US/07866979
Patent No. 5532347
GENERAL INFORMATION:
APPLICANT: Cone, Roger D
APPLICANT: Mountjoy, Kathleen G
TITLE OF INVENTION: Melanocyte Stimulating Hormone Receptor
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegritti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/866,979
FILING DATE: 19920410
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532347nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 92,154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 15..959
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..14
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 960..1260
US-07-866-979-3

Query Match 29.1%: Score 29.4; DB 1; Length 1260;
Best Local Similarity 58.6%: Pred. No. 0.33;
Matches 51: Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 2 ctcaagttcctgacacatcttggtccctgcgcgcgcctctctctacgcact 61
DB 363 CTCATGACGTGCTCATCTGTGCTCCATGCTGTCAGTCTGCTTCGCGCATCATTT 422
OY 62 atccaagattgggcatcgcgcacatccgcagactgac 88


```

1  APPLICANT:  Cong, Roger D
2  APPLICANT:  Mounjoy, Kathleen G
3  TITLE OF INVENTION:  Melanocyte Stimulating Hormone Receptor
4  TITLE OF INVENTION:  and Uses
5  NUMBER OF SEQUENCES:  8
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE:  McDonnell
8  STREET:  300 South Wacker Drive
9  CITY:  Chicago
10 STATE:  IL
11 COUNTRY:  USA
12 ZIP:  60606
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE:  Floppy disk
15 COMPUTER:  IBM PC compatible
16 OPERATING SYSTEM:  PC-DOS/MS-DOS
17 SOFTWARE:  Patent In Release #1.0, Version #1.25
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/09/201,746
20 FILING DATE:  01-DEC-1998
21 CLASSIFICATION:  435
22 ATTORNEY/AGENT INFORMATION:
23 NAME:  No. 626821nan, Kevin E
24 REGISTRATION NUMBER:  35,303
25 REFERENCE/DOCKET NUMBER:  92,154-J
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE:  312-913-0001
28 TELEFAX:  312-913-0002
29 TELEX:
30 INFORMATION FOR SEQ ID NO:  3:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH:  1260 base pairs
33 TYPE:  nucleic acid
34 STRANDEDNESS:  single
35 TOPOLOGY:  linear
36 MOLECULE TYPE:  cDNA to mRNA
37 FEATURE:
38 NAME/KEY:  5'UTR
39 LOCATION:  1..14
40 FEATURE:
41 NAME/KEY:  CDS
42 LOCATION:  15..959
43 FEATURE:
44 NAME/KEY:  3'UTR
45 LOCATION:  960..1260
46
47 US-09-201-746-3
48
49 Query Match          29.1%; Score 29.4; DB 4; Length 1260;
50 Best Local Similarity 58.6%; Pred. No. 0.33;
51 Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0.
52
53 QY      2  ctcaagcttcgcgcacatcttggctccctgctgcccgcggccctcctctaccgcact 61
54          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
55 Db      363  ctcatgtacgcgtgcacatctgtgcctccatgctgcctcactgctcctctgctgcatcatt 422
56
57 QY      62  atccaagatcggcgcacatc 88
58          |||  ||||  ||||  ||||
59 Db      423  gctatagacgcgtacatctccatcttc 449
60
61 RESULT      7
62 US-09-097-231-3
63 ; Sequence 3, Application US/09097231
64 ; Patent No. 6278038
65 ;
66 GENERAL INFORMATION:
67 ;
68 APPLICANT:  Cong, Roger D
69              Chen, Wenbiao
70              Low, Malcolm J
71 TITLE OF INVENTION:  Mammalian Melanocortin Receptor and Uses
72 NUMBER OF SEQUENCES:  22
73 CORRESPONDENCE ADDRESS:
74 ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff

```

```

1 STREET: 300 South Wacker Drive
2 CITY: Chicago
3 STATE: Illinois
4 COUNTRY: USA
5 ZIP: 60606
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: Patentin Release #1.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/09/097,231
15 FILING DATE: 12-Jun-1998
16 CLASSIFICATION: <Unknown>
17 ATTORNEY/AGENT INFORMATION:
18 NAME: No. 627803Boan, Kevin E
19 REGISTRATION NUMBER: 35,703
20 REFERENCE/DOCKET NUMBER: 86,86-C
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: 312-913-0001
23 TELEFAX: 312-913-0002
24 TELEX: <Unknown>
25
26 INFORMATION FOR SEQ ID NO: 3:
27
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 1260 base pairs
30 TYPE: nucleic acid
31 STRANDEDNESS: single
32 TOPOLOGY: linear
33 MOLECULE TYPE: cDNA to mRNA
34
35 FEATURE:
36 NAME/KEY: 5'UTR
37 LOCATION: 1..14
38
39 FEATURE:
40 NAME/KEY: CDS
41 LOCATION: 15..959
42
43 FEATURE:
44 NAME/KEY: 3'UTR
45 LOCATION: 960..1260
46
47 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
48
49 US-09-097-231-3
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
```

```

: COUNTRY: USA
: ZIP: 07065-0900
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/759,848
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/281,393
: FILING DATE: 27-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Wallen III, John W.
: REGISTRATION NUMBER: 35,403
: REFERENCE/DOCKET NUMBER: 19234
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 594-4720
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1378 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-759-848-2

Query Match      27.5%; Score 27.8; DB 1; Length 1378;
Best Local Similarity 65.1%; Pred. No. 1.2;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 4 caagtcctgcacatcttggtccctgctgcgccgcccctctctctacccgcaact 63
   ||| ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
DB 816 CATGCTCCTGAATGCTGCGCTTCTGCTGCCCTGAGTGTATCACCCTTTCGACGAT 875
   ||| ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
QY 64 cca 66
   ||
DB 876 GCA 878

RESULT 9
PCT-US95-09383-2
: Sequence 2, Application PC/TUS9509383
: GENERAL INFORMATION:
: APPLICANT: Borkowski, Joseph A.
: APPLICANT: Strader, Catherine D.
: APPLICANT: Hess, John W.
: APPLICANT: Chen, Howard Y.
: APPLICANT: Trumbauer, Myrna E.
: TITLE OF INVENTION: BRADYKININ B2 RECEPTOR MODIFIED
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: John W. Wallen III
: STREET: 126 E. Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07065-0900
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09383
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
```

```

: APPLICATION NUMBER: US 08/281,393
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wallen III, John W.
: REGISTRATION NUMBER: 35,403
: REFERENCE/DOCKET NUMBER: 19234
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 594-4720
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1378 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: PCT-US95-09383-2

Query Match      27.5%; Score 27.8; DB 5; Length 1378;
Best Local Similarity 65.1%; Pred. No. 1.2;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 4 caagtcctgcacatcttggtccctgctgcgccgcccctctctctacccgcaact 63
   ||| ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
DB 816 CATGCTCCTGAATGCTGCGCTTCTGCTGCCCTGAGTGTATCACCCTTTCGACGAT 875
   ||| ||||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
QY 64 cca 66
   ||
DB 876 GCA 878

RESULT 10
US-08-938-291A-3/C
: Sequence 3, Application US/08938291A
: Patent No. 6117673
: GENERAL INFORMATION:
: APPLICANT: Lev, Sima
: APPLICANT: Plozman, Gregory D.
: APPLICANT: Schlessinger, Joseph
: TITLE OF INVENTION: RDBG PROTEINS AND RELATED
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FastSeq
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/938,291A
: FILING DATE: September 26, 1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/027,337
: FILING DATE: October 11, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Wardburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 228/172
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
```



```
; Sequence 16, Application US/08479285
; Patent No. 6207815
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P O BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,285
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040687
; FILING DATE: 31-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: ULMER, DUANE C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-37,075C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-479-285-16

Query Match 25.3%; Score 25.6; DB 4: Length 1984;
Best Local Similarity 59.7%; Pred. No. 7.9;
Matches 43: Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 2 ctcaagttcctgacacatcttgctgcctgcctgcctctctctacccgact 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 897 ctgagatccctacacagattttttcccaacttctctcagccgctcctcagggcaaat 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 62 atccaagattcg 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 837 atccaagattag 826

RESULT 14
US-08-481-337A-1/C
; Sequence 1, Application US/08481337A
; Patent No. 5863738
; GENERAL INFORMATION:
; APPLICANT: TEN DIJKE, Peter
; APPLICANT: HELDIN, Carl-Henrik
; APPLICANT: MIYAZONO, Kohel
; APPLICANT: SAMPATH, Kuber T.
; TITLE OF INVENTION: Morphogenic Protein-Specific Cell
; TITLE OF INVENTION: Surface Receptors and Uses Therefor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
```

```
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,337A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-097CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1509 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1509
; OTHER INFORMATION: /product="Human ALK1"
; US-08-481-337A-1

Query Match 25.1%; Score 25.4; DB 2: Length 1509;
Best Local Similarity 64.4%; Pred. No. 8.5;
Matches 38: Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 4 caagttccctgcacacatcttgctgcctgcctgcctgcctctctacccgact 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 216 caagttccctgcacacagcttctctgagggctcctcctcctccagcagcacta 158

RESULT 15
US-08-696-268B-1/C
; Sequence 1, Application US/08696268B
; Patent No. 5968752
; GENERAL INFORMATION:
; APPLICANT: ICHIO, HIDENORI
; APPLICANT: NISHITO, HIDEKI
; APPLICANT: SAMPATH, KUBER T.
; TITLE OF INVENTION: NOVEL SIGNALING RECEPTOR FOR
; TITLE OF INVENTION: MORPHOGENIC PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,268B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, Thomas C.
```



```

:   REGISTRATION NUMBER: 36,989
:   REFERENCE/DOCKET NUMBER: CRP-117
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (617) 248-7000
:   TELEFAX: (617) 248-7100
:   INFORMATION FOR SEQ ID NO: 1:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 1509 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA
:   FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 1..1509
:   OTHER INFORMATION: /product="Human ALK-1"
:
US-08-696-268B-1
    
```

```

Query Match      25.1% Score 25.4; DB 2; Length 1509;
Best Local Similarity 64.4% Pred. No. 8.5;
Matches 38: Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      4 caagtcctgcacacatcttggtccctgctgcgccgccccctcctcctctacgcgacta 62
        1111111111111111111111111111111111111111111111111111111111111111
DB      216 CAAGTTCCCGCAGCCCGATGTTCTGTGGGGGTGCTCCCTCCCTCCCGCAGCAGACTA 158
    
```

Search completed: October 3, 2002, 16:24:54
 Job time: 17134 sec